7.00 per volume
Single numbers $4.00

QUARTERLY

Two volumes per year

April, 1934
Volume X, No. 2

Founded in 1927 by Edward Bradford Titchener and Carl Murchison

THE JOURNAL OF GENERAL PSYCHOLOGY

Experimental, Theoretical, Clinical, and Historical Psychology

APRIL, 1934

Chaque article est suivi d'un résumé en français
Jedem Artikel wird ein Referat auf deutsch folgen

THE EFFECT OF PRACTICE ON INDIVIDUAL DIFFERENCES WHEN STUDIED WITH MEASUREMENTS WEIGHTED FOR DIFFICULTY 249
By Harry Ewert

THE ALGEBRA OF LUMINOSITY IN RELATION TO WHITE COMPLEMENTARY PAIRS OF WAVE-LENGTHS . . 286
By R. D. Williams

PSYCHOPHYSICS AND BEHAVIOR 299
By C. H. Graham

SPATIAL SUMMATION OF SUBLIMINAL STIMULI IN THE RETINA OF THE HUMAN EYE 311
By Robert J. Beitel, Jr.

A STATISTICAL STUDY OF THE RELATIONS OF NEURASTHENIC, DYSPEPTIC, AND ALLERGIC SYMPTOMS . 328
By Carlos Kling

INDIVIDUAL DIFFERENCES IN ABILITY TO ALTERNATE ACTIVITIES 344
By Clara Burri

PSYCHOLOGY AND TECHNIQUES OF DISCOVERY . . . 364
By T. N. Whitehead

THE INFLUENCE OF SYMBOLIC REHEARSAL UPON THE RETENTION OF A MAZE HABIT 376
By Robert S. Sackett

(over)

Worcester, Massachusetts
Copyright, 1933, by Clark University
Entered as second-class matter December 1, 1927, at the post-office at Worcester, Mass., under the Act of March 3, 1879

AN ATTEMPT TO MEASURE CERTAIN PHASES OF SPEECH 399
By Joseph W. Hawthorne

PSYCHOPHYSIOLOGICAL MEASUREMENTS DURING SOMNOLENCE INDUCED BY PHENOBARBITAL (LUMINAL) 415
By Louis H. Cohen, Robert R. Sears, Stanley B. Lindley, Walter C. Shipley, and Pauline K. Snedden

A STUDY OF IMAGINED COLOR BLENDS 432
By Kate Gordon

SHORT ARTICLES AND NOTES 440

An Investigation of the Mechanism of Caloric Nystagmus in the Rabbit (Joseph Gidansky) 440

On the Structure of Acts (Donald K. Adams and T. L. McCulloch) 450

Perseverative Tendency in Motor Activity (K. H. Rogers) . . 455

An Explanation of Metfessel's Stroboscopic Effect Observed in Mirror Vibration (Neil Warren) 463

Influence of the Speed Attitude on Test Performance (Miles A. Tinker) 465

Widespread Improvement in Tactile Discrimination Following Localized Training (Shepherd Ivory Franz and Sybil Kilduff) 469

The Effect of Combining Some Geometrical Optical Illusions (Chung-Fang Ni) 472

APPARATUS 477

An Examination of Electrical-Current–Stabilizing Devices for Psychological Experiments (Karl F. Muenzinger and Frank C. Walz) 477

A Universal Flexible Elevated Finger Maze (Ralph G. Whisler) 482

An Electrical Maze Stylus (Ralph W. Gilbert) 485

BOOKS 487

Books Received 487

THE EFFECT OF PRACTICE ON INDIVIDUAL DIFFERENCES WHEN STUDIED WITH MEASUREMENTS WEIGHTED FOR DIFFICULTY*

From the Department of Psychology of the University of Vermont

HARRY EWERT

I. INTRODUCTION[1]

This investigation endeavors to develop a statistical technique which takes account of several urgent requirements that have hitherto not received adequate consideration in studies on the effect of practice upon individual differences. Hunter (16) has pointed out that we are not justified in accepting as final the diverse conclusions reached by experimenters working in this field. He says:

> As Chapman [5] and Stoddard [28] in particular have noted, certain factors which are vital and indispensable for a solution of the problem have been neglected by experimenters. (1) The subjects used have not only usually been too few in number but they have not been selected with reference to learning ability nor can we be reasonably sure that adequate incentives have been present, particularly in the final stages of training where improvement is most difficult. . . (2) Account must be taken of the arbitrary character of the scale used for scoring . . . In the absence of a determination of a true zero, no method of determining individual differences in terms of score ratios can properly be employed. . . . (3) . . . the third important factor . . . is that of the inequality of units of achievement. . . . Ordinarily the units of achievement in the early part of training are easier than those in the middle portion of the learning process, and the final units are the most difficult. . . . As the limits of achievement are reached, and particularly as a performance approaches a physiological limit, the additional

*Recommended by Walter S. Hunter and accepted for publication by Carl Murchison of the Editorial Board.

[1]This research was begun in the autumn of 1931. A few of the major findings of the present paper were reported at the 1933 meeting of the American Psychological Association at Chicago.

The writer is deeply indebted to W. S. Hunter for reading and criticizing the manuscript.

> increments of accomplishment become more and more difficult to attain, and fewer and fewer subjects can master them. It therefore follows that differences in individual performance may actually increase although the gross scores may indicate a decrease. (pp. 616-619)

I am quoting Hunter at length to show that the underlying factors which may lead to the eventual solution of this problem have been recognized and stated by him and others. However, there has not appeared to my knowledge any attempt to bring these factors under experimental control.

The immediate purpose of this investigation is to study the effect of practice upon individual differences (1) when units of measurement are weighted in terms of difficulty, (2) when they are based upon a scale on which scores have a common zero, and (3) when they are expressed in ratios of like function. Units of measurement in learning have commonly been expressed in reciprocals of unlike ratios such as time per work or work per time, in which case time or work is kept constant and achievement is varied. This leads to diverse interpretation of results as has been noted by Peterson (21), Reed (24, 25), and others. In order to overcome these difficulties I have expressed units of measurement in terms of like functions such as average amount of individual time (secs.) necessary to eliminate a unit of the average amount of total time (secs.) of practice, or, similarly, the amount of work (errors), necessary to eliminate a unit of work. Under these conditions achievement remains constant and time and work become the variables. Difficulty is then expressed in terms of the number of time and work units necessary to progress from one achievement score to another.

II. Historical Background of Problem

There is considerable disagreement in previous investigations concerning the effect of practice upon individual differences. The early investigators (and some recent ones) (29, 30, 31, 6, 12, 7, 1, 13, 3, 4, 33, 23, etc.) report divergence of individual ability as a result of practice. However, a preponderance of the more recent experimenters (11, 24, 25, 18, 22, 20, etc.) report that individual ability converges as a result of practice. There are also those (14, 8, 17, etc.) who observe that individual ability remains constant with practice.

In the matter of methodology there is no more agreement than in the conclusions reached. A host of methods for the computation of data have been employed with perplexing results. Some appear to show divergence, others convergence, and still others that practice does not change individual differences. I shall now review some of the conditions or factors which have been deemed responsible for the diversity of results secured with the different methods.

1. *Units of Measurement.* Wells (39), working with addition and number-checking tests, found that curves plotted on absolute gain in amount of work per time diverged. He, however, pointed out that, if plotted on the basis of time per unit of work, the curves would converge. Peterson (21), criticizing Thurstone's (33) experiment, pointed out more specifically that individual differences, as a result of practice, cannot be determined upon scores of absolute units such as work per time or time per work since these two conditions yield different results. Treating Thurstone's data on the basis of time per unit of work he found convergence, whereas Thurstone, who had computed his own data in terms of work per time, had found a divergence. Reed (25) in 1931 says of his earlier work:

> Reed in 1924 pointed out that the conclusion of leading psychologists that practice increased individual differences was false because of the measures on which it was based. He said that his conclusion was justified when gains are measured by absolute increases in amount of work per unit of time but that an entirely opposite conclusion is reached when gains are measured by time per unit of work, and that because of this contradiction, neither measure is valid. The difficulty with either method is that psychologists have been handling them like whole numbers, but in reality they are fractions such as $\frac{G}{T}$ or $\frac{T}{G}$ where T and G represent time and gain respectively, and that fractions cannot be compared by dealing only with the numerators. (p. 13)

Chapman (5) adequately criticized Reed's ratio method on the ground that, since Reed's scores were made on an arbitrary scale with an arbitrary zero point, small amounts of practice would by the ratio method assure infinite amounts of improvement. To this Reed (25) replies:

> . . . it may be stated that the ratio method, just like every other

> method, would increase in validity if used on a scale having an absolute zero, but when applied to scores of an arbitrary scale, it is just as valid as other measures of this sort. It is also true that the method cannot be used on zero scores but that does not invalidate its use with scores that are greater than zero. (p. 14)

Besides these limitations recognized by Reed himself another difficulty with his ratio method is that he does not tell us which of the two ratios should be the numerator and which the denominator.

This brief historical summary concerning units of measurement establishes the following salient points: (*a*) Methods of computation when based upon amount of work per unit of time and amount of time per unit of work give different results. (*b*) Neither method can be said to be better since all depends upon the formulation of the problem. (*c*) The ratio method is not valid unless based upon a scale having an absolute zero. To this I should like to add that any method which correlates time with work or work with time (which might be called the "achievement method") is in itself not adequate to express relative variability since it does not take into account the changing difficulty values at successive stages of practice.

2. *Validity of Methodology and Conclusions Concerning Individual Differences as Based upon the Amount of Support Received from Authors and Data at Large.* Some investigators, notably Kincaid (18) and Reed (25), have surveyed the literature with the purpose of testing the various methods and conclusions by the amount of support they have received from authors and existing data at large. Although such a survey yields interesting results and aids in analysis of technique, it furnishes no ultimate proof of the validity of a method since different methods may be applied to the same data with diverse results. And as for conclusions we must of necessity rely upon the validity of the methods employed.

3. *Complexity of Material.* A group of experimenters have attributed the diversity of opinion concerning the effect of practice upon individual differences to differences in the complexity of the test material used. Barlow (2) found that continued practice in card-sorting resulted in a decrease of individual differences while substitution tests and mental multiplication showed an increase. He concluded (1) that the more complex the trait learned, the

greater the divergence of scores with equal amounts of practice and (2) that the simpler the task, the greater is the convergence. In a later paper by Peterson and Barlow (22) similar conclusions are drawn. These authors, however, suggest the need of further work on this phase of the problem. Ruch (26), correlating Army Alpha scores and achievement, found that practice caused individual differences to increase in abstract relations, decrease in card-sorting, and remain constant in code substitution. However, in so far as these correlations are drawn between different types of material, their true significance in the same material remains unknown. In a recent study Mattson (19) investigated the effect of complexity of material upon the form of the learning curve. Three maze patterns of different complexity were used. Errors were automatically recorded. She concludes:

> Coefficients of time and error scores seem to support the trend noted for variability; as practice continues they [coefficients] decrease upon the simple task, remain practically unchanged on intermediate tasks, and increase upon the complex task. (p. 380)

It should, however, be pointed out that these conclusions hold only for time and error coefficients. When she computes relative and absolute gain, she finds negative correlation with initial status. But, strangely enough, she regards this as representing a positive relationship: "In other words, those who start out with a low score in time and errors, indicative of high maze performance, also gained the most" (p. 381). If "high maze performance" is defined in terms of relative gain then the relation may be regarded as positive. If, however, final performance be regarded as "high maze performance," the opposite would be true. I shall present empirical evidence in this paper to show that, unless scores are weighted for difficulty, relative gain is not indicative of relative improvement.

4. *Individual Differences as Dependent upon Factors Inherent in Methodology.* A number of investigators have suggested that whether abilities diverge or converge as a result of practice depends upon such factors as (*a*) the criteria set for learning (25), (*b*) the method of computing the data (25, 21, 18), (*c*) the units of measurement chosen (25, 21), (*d*) the extent of practice (4), (*e*) the type of material under consideration (19), (*f*) the degree of motivation (5), and especially (*g*) the complexity of the material

(2, 19). In this connection it might be said that, whereas these factors will undoubtedly influence the *amount* of variation, yet they are not to be taken as basic determiners of individual differences since they are merely factors in a measuring technique and not the phenomena measured. In other words, these are the factors to be controlled in studies of individual differences rather than to be regarded as being themselves responsible for such differences. The thesis which I wish to develop in this paper is *that, when certain basic principles which have hitherto not been taken into account are carefully considered, it will be possible to establish more agreement concerning the effect of practice upon individual differences; and, furthermore, that this will be accomplished without a mere enumeration of the environmental factors involved.*

III. Experimental Section

Experiment I

Kincaid's (18) and Reed's (25) analyses of the experimental data on individual differences have furnished ample proof that different results may be derived from the same data when different methods of computation are employed. The purpose of this preliminary experiment is to reexamine some of the accepted methods of computation, to describe briefly a method which has not been previously applied to this problem, and, if possible, to show some reasons why various methods reported in the literature appear to give different results.

Test Material and Subjects. A letter-digit substitution test described in Foster's *Experiments in Psychology, Teacher's Manual,* revised by Tinker (37), was given to 94 college students (sophomores, juniors, and seniors), according to the directions of the manual. Each subject received 20 one-minute trials.

Motivation. The instructions to the subjects in Tinker's manual were supplemented by offering a first prize of $2.00 and a second prize of $1.00 in reward for the best score (number of substitutions minus errors).

Treatment of Data and Results. The data are graphically summarized in the frequency distribution curves of Figure 1. On the abscissa is represented the achievement range. On the ordinate is represented the frequency of achievement scores. The unbroken

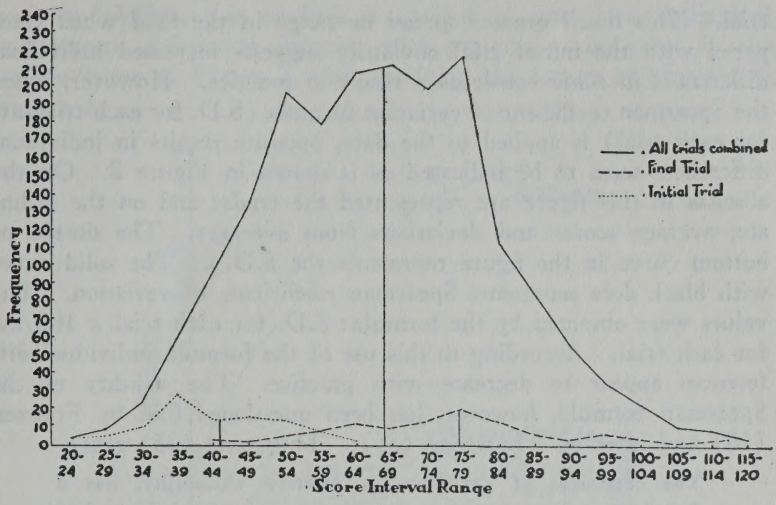


FIGURE 1

FREQUENCY DISTRIBUTIONS OF ACHIEVEMENT SCORES DURING INITIAL AND FINAL TRIALS IN LETTER-DIGIT SUBSTITUTION AS COMPARED WITH DISTRIBUTION OF ACHIEVEMENT IN ALL TRIALS COMBINED

On the abscissa is represented the total achievement range and on the ordinate the frequency of achievement scores. Frequency of achievement is considered irrespective of the trials into which it falls.

line indicates the distribution of performance for the whole group in the entire practice series. Thus there were 3 scores made in the entire practice series which ranged between 20-24, 6 scores between 25-29, 23 between 30-34, etc. It will be seen that achievement scores are here pictured irrespective of the trials into which they fall. The distribution thus gives us, at a glance, a complete picture of the frequency of achievement scores made by the group in 20 one-minute practice trials. It is now possible to present pictorially the relationship of initial and final performance to total performance. The dotted line to the left represents the distribution of performance in the first trial; and the broken line to the right, that for the final trial. The vertical lines indicate the means of the distributions. It will be seen that in the initial trial the best score does not reach the mean for total achievement, whereas in the final trial the poorest score overlaps the mean of the first

trial. This much greater spread or range in the final when compared with the initial trial obviously suggests increased individual differences in achievement as a result of practice. However, when the Spearman coefficient of variation formula (S.D. for each trial/av. for each trial) is applied to the data, opposite results in individual differences seem to be indicated as is shown in Figure 2. On the abscissa in this figure are represented the trials; and on the ordinate, average scores and deviations from averages. The dotted or bottom curve in the figure represents the S.D.'s. The solid curve with black dots represents Spearman coefficients of variation. The values were obtained by the formula: S.D. for each trial x 100/av. for each trial. According to this use of the formula individual differences appear to decrease with practice. The validity of the Spearman formula, however, has been questioned first by Franzen (10) and later by Thurstone (35). Thurstone (35) says:

> The coefficient of variation, or relative variability, has a definite meaning when the measurements to which it is applied are made from a true origin but the coefficient of variation is sheer nonsense when it is applied to measurements from an arbitrary origin. (pp. 188-189)

If, in the face of such criticism, we wish to determine the relationship between the S.D.'s and the decreasing or increasing averages in successive trials, it would seem more logical to divide the S.D.'s by a constant instead of by other variables, such as averages of respective trials. In the field of mental testing we have determined norms, constants, or levels of ability from the actual performance of a cross section of society. Comparison of ability is then made on the basis of such constants. If this method were applied to the total distribution of a practice series the average of the entire series might represent the constant for such a distribution of scores. This total average could then be treated as a constant for a given group engaged in a given activity in an allotted practice period. Under these conditions the ratios between the S.D.'s and the changing values of scores during practice are expressed as S.D. for each trial/*total* av., instead of in the conventional manner, namely, S.D. for each trial/av. for *each trial*. The former expression satisfies the very purpose for which the coefficient of variation was devised; but, instead of involving two variables, it furnishes one constant which serves as a check on increasing or decreasing scores during

practice. When this method is applied to our data, individual achievement differences increase with practice as indicated in the broken curve of Figure 2, thus verifying the results as pictured in Figure 1.

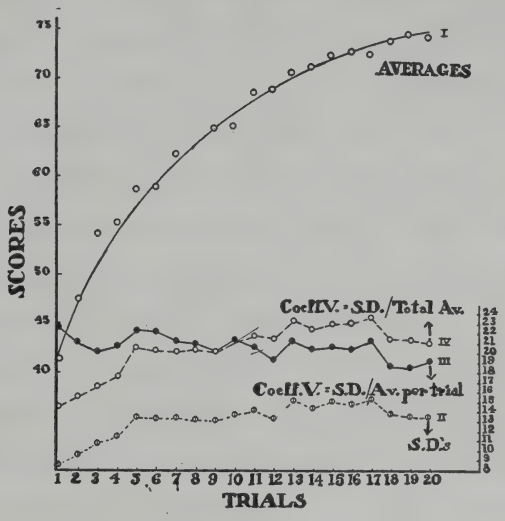


FIGURE 2

AVERAGES, STANDARD DEVIATIONS, AND COEFFICIENTS OF VARIATION OF LETTER-DIGIT SUBSTITUTION COMPUTED FOR EACH TRIAL

On the abscissa are represented trials. The figures on the left ordinate indicate average achievement scores, while those on the right ordinate represent a magnified scale for S.D.'s and coefficients of variation in achievement. Curve I gives average achievement per trial and curve II gives the S.D.'s of these averages. Curve III gives coefficients of variation obtained by the formula:

S.D. x 100 for each trial/average for *each trial*

Curve IV gives coefficients obtained by the formula:

S.D. x 100 for each trial/total average

Table 1 gives coefficients of correlation which were derived by the use of the rank-difference method. It will be seen that the

TABLE 1

Initial and final scores	.75±.031
Initial scores and intelligence (raw scores)*	.67±.067
Final scores and intelligence	.34±.070
Initial scores and percentage of gain	—.50±.054
Percentage of gain and intelligence	—.18±.090

*Intelligence scores were those on the Thurstone and Thurstone Psychological Examination which is annually given to all freshmen at the University of Vermont.

first three correlations are positive and the last two are negative. The first correlation signifies a tendency for subjects to maintain their respective ranks from initial to final performance. The second and third show a positive relationship between initial performance and intelligence and between final performance and intelligence, respectively. There is thus a positive relationship between achievement and intelligence rank. However, when we compare subjects as to their relative range of achievement, we detect an inverse relationship as is indicated in the fourth correlation of Table 1. There is likewise a negative relation between intelligence and relative achievement as shown in the last correlation of the table. This means that the higher the initial performance the smaller the relative range of progress in achievement appears to be, and subjects with a high intelligence score appear to progress relatively less than subjects with lower intelligence scores. Although these findings seem paradoxical, similar findings have previously been reported for motor activity (25, 18, 19, 15, etc.). The question now arises, shall we accept the derived negative relationship between initial performance and relative gain or intelligence and relative gain as finally established? The answer to this question awaits a definition of some of the terms commonly used in connection with the learning process and a more careful analysis of the changing values of our units of measurement as learning approaches the physiological limit. Data will be presented subsequently to show that relative improvement is inextricably related to the changing values of units of measurement, whereas gross gain in achievement may be considered irrespective of such changing values, because the relative difficulty values cancel each other when initial and final scores are subtracted one from another.

At this point I should like to draw a distinction between the terms "improvement" and "achievement" as related to practice. Improvement is characterized by the amount of work necessary to reach a certain point in a given activity, whereas achievement refers to the point reached. Improvement is thus a change of trend from worse to better or more to less in achievement, which may require much or little work. When we compare individual differences of initial and final scores we are dealing with differences in achievement, whereas when we refer to individual differences in amount of work necessary to progress from one achievement score to another we are

characterizing differences in improvement. Consequently, when achievement scores are treated independently of improvement the computation should be in terms of gross differences since relative scores always involve the difficulty factor. These definitions are based upon the concept that improvement consists in elimination of errors or excessive movements and of reduction of time. If we accept the above terminology, we can say, in accordance with the above findings, that there is a positive relationship between initial and final achievement. And there appear to be indications that individual differences in *achievement* increase with practice. But does it also follow from the above findings that there is a negative relationship between relative initial and final *improvement?* We remain in total ignorance about improvement as such until we know what happens between the various achievement scores, and until we are able to measure the changing difficulty values of work units as a result of an approach to the physiological limit. What happens between achievement scores cannot be determined by expressing measurements in terms of unlike ratios such as work per time or time per work, since by this method we cannot determine the exact change of time or the exact change of work but only a change in their ratios. It is thus necessary to express units of measurement in terms of like functions such as work per work or time per time. By this method it is possible to determine the number of time units necessary to eliminate a time unit (a second, for example) or the number of work units necessary to eliminate a work unit (such as an error). This method enables us to determine how many time or work units exist between achievement scores. *Consequently, it establishes the difficulty value for achievement scores.* This method (as will be explained more fully below) further requires that in order to compute group averages the achievement scores for the members of the group must overlap in size, or at least fall within a common interval. It thus avoids the bugbear common to attempts of comparing unequal initial and final scores. By the use of such a method it is possible to make direct comparison of the various subjects in a group since they work within common levels of difficulty. Hence it avoids the complications inherent in conventional methods which attempt to measure improvement by treating units of measurement independently of unequal individual physiological zeros. Since the method furnishes some information as to what

occurs in successive increments and decrements of learning, it might be called the "improvement method" to distinguish it from the ratio method which we might call the "achievement method." We shall next describe an experiment in which the "improvement method" has been applied.

Experiment II

As stated in the introduction, the purpose of this experiment is to study the effect of practice upon individual differences when units of measurement are weighted in terms of difficulty and when they are expressed in terms of like function. In order to accomplish this, the following requirements are essential:

1. The type of the problem should be one (*a*) which permits a wide range of maximal and minimal achievement; (*b*) which is novel to the subjects; and (*c*) in which units of measurement can be expressed in terms of like function.

2. Practice should be prolonged in order to obtain a wide range of achievement scores which overlap in size for all individuals in the group. The most fruitful practice level for comparison of individual differences is the achievement range over which all of the subjects progress during the course of practice. The achievement-score levels which do not overlap in size for all subjects, commonly found at the beginning and end of practice, present difficulties which are hard to overcome. (*a*) The units of measurement are so large or small (depending upon whether time or work units are used) during initial and final performance that a given technique loses its sensitiveness. (*b*) Since the achievement scores for different subjects in a group do not overlap during initial and final performance, their difficulty must be based upon the performance of an increasing (initial performance) or decreasing (final performance) number of subjects.

3. The method of measurement should be one (*a*) which is sensitive to slight individual differences; (*b*) which expresses units of measurement in terms of like function; (*c*) which takes account of individual levels of difficulty, and thus (*d*) furnishes a common zero for the different individual rates of deceleration in achievement as the various individuals approach their physiological limits.

4. Motivation should be highly constant and should encourage the highest individual achievement possible.

Apparatus. The apparatus used was a modified Snoddy (27) mirror-tracing device. The apparatus contains a star-shaped groove, 8 mm. in width, cut in a copper plate. The groove is traced by means of a metal stylus under mirror-drawing conditions. The stylus is attached to an electric counter by means of a cord. Errors are recorded electrically every time the stylus touches either side of the groove. An ink-spot on the star indicates where the subject is to start and stop tracing.

Procedure. Each subject received 100 trials which required from 2 to 2½ hours per subject. All tests were conducted individually. A single tracing of the star constituted a trial. After each five successive trials a one-minute rest period was allowed. Time and error records were recorded on a mimeographed sheet which had been previously prepared. The star was traced clockwise.

Motivation and Instructions. Two types of motivation were used with all subjects to insure maximum and constant effort. The first consisted of a cash prize of $5.00 awarded to the subject making the best "time plus error" score. The second consisted of a penny scheme. At the beginning of the experiment about 100 pennies were placed upon the table. Then S was instructed as follows:

> You will have an opportunity to win a $5.00 prize by making the best score in the group of subjects who take part in this experiment. By best score we mean one which is lowest in time and errors. So you will have to work for both accuracy and speed. You will also have the opportunity to win as many of the pennies on the table as you can by observing the following rules: After reaching a time plus error score of 40 you will be given 5 pennies to start with. You will win an additional penny for every two points of improvement until you reach a score of 25, after which you will win a penny a point. On the other hand, if you make a worse score by three points after reaching 40 or a worse score by two points after reaching 25, you will lose a penny for each such recession.

Every time a penny was earned it was immediately added to the subject's original award, and every time a penny was lost it was immediately subtracted from his "capital." Naturally the subjects were greatly interested in a growing "capital" and they worked hard to retain their possessions.

Subjects. Twenty women and 20 men served in the experi-

ment. They were college sophomores, juniors, and seniors. All subjects were right-handed. They all testified to lack of experience in mirror-tracing.

Results. The correlation between initial achievement and gross gain in achievement was found to be .98±.008. This unusually high correlation is probably a result of the type of motivation used in this experiment. The subjects did not dare to fluctuate much in their performance for fear of losing pennies which they had previously earned. It thus appears from the results that the type of motivation used insures a high constancy of effort throughout the practice period.

The steps taken in order to express units of measurement in terms of a like function may be described as follows:

1. Find the achievement range (in time or errors) for the whole group in the entire practice series. This will be the difference between the worst score in the initial trial and the best score in the final trial. It may be illustrated by the abscissa in Figure 1, where the range is 20-120, or it may be illustrated in Table 3, where the range for time in seconds is given in the interval column as 245-4.

2. Such a range divides itself into three natural intervals. The middle interval consists of the respective achievement scores over which all subjects progress during the course of practice. The end intervals consist of achievement scores whose sizes do not overlap. The first-end interval includes the range of scores up to the point where the best subject starts the performance and the second-end interval includes the range between the points where the poorest and the best subject drops out. For time in this study these intervals are: 245 (seconds) to 45, first-end interval; 45 to 10, middle interval; 10 to 4, last-end interval. For errors they are: 200 to 25, first-end interval; 25 to 10, middle interval; 10 to 0, last interval (see Table 3).

3. It is further possible to subdivide the three natural intervals into smaller intervals. The choice of the intervals may be arbitrary, providing scores or score averages are ultimately reduced to a common basis by dividing them by the chosen step intervals. For time in seconds the intervals chosen in this study are, 245-145, 145-95, 95-45 for the first-end interval; then steps of 5 from 45-10 for the middle or overlapping interval; and, finally, steps of 1 from 10-4 for the last-end interval (see Table 3). It was believed that by

TABLE 2
THE RAW TIME SCORES IN SECONDS FOR TWO OF THE SUBJECTS

	Subject I				Subject II		
Trial	Time in secs.	Trial	Time in secs.	Trial	Time in secs.	Trial	Time in secs.
1	105	51	13	1	219	51	14
2	**103**	52	11.6	2	**121**	52	15
3	**39**	53	13.4	3	111	53	14.5
4	37	54	14	4	**68**	54	14
5	41	55	13.5	5	54	55	12
6	44	56	12	6	56	56	13.8
7	**24**	57	11.5	7	47	57	12
8	24	58	12	8	**30**	58	13.8
9	25	59	11	9	32	59	13.4
10	22	60	13	10	31	60	12.5
11	26	61	13.3	11	**24**	61	13
12	**20**	62	13	12	30	62	12
13	18	63	12	13	31	63	11.6
14	17	64	11.8	14	**20**	64	11.4
15	16	65	11	15	20	65	**9.5**
16	19	66	12	16	22	66	11.8
17	**15**	67	13	17	23	67	10.2
18	16.5	68	12	18	21	68	9.8
19	15.5	69	13.6	19	23	69	11.5
20	15	70	12	20	23	70	13
21	18.5	71	11.5	21	20	71	11.4
22	20	72	11.4	22	21	72	12
23	16	73	11.5	23	19	73	11
24	13	74	11.5	24	20	74	12
25	14	75	**10**	25	18	75	12
26	15.5	76	9	26	18	76	11.5
27	15	77	9.5	27	19	77	11.5
28	18.5	78	8.5	28	16	78	10
29	14	79	8.4	29	**15**	79	10
30	15.5	80	9.5	30	15	80	11
31	16	81	8.2	31	16	81	11
32	18.5	82	10	32	18	82	10
33	14	83	9	33	16	83	9.5
34	14	84	11	34	16	84	10
35	16	85	9.5	35	15	85	11
36	14.5	86	11.5	36	18	86	10.5
37	15.5	87	10.8	37	19	87	10.2
38	15	88	10.4	38	16	88	10
39	16	89	9.5	39	15	89	10
40	18	90	**8**	40	13.5	90	10
41	18	91	7.5	41	14	91	**9**
42	16	92	8.5	42	14	92	10
43	15	93	9	43	13.5	93	10
44	13	94	9.6	44	12	94	9.5
45	13.4	95	9	45	12	95	10
46	12.5	96	8.5	46	15	96	9.4
47	13.5	97	**7**	47	12.8	97	9.5
48	14	98	10	48	13	98	9.5
49	15	99	10	49	15	99	9.2
50	14.3	100	9	50	14	100	9.2

tapering off the achievement scale in this manner it would be possible to obtain a more uniform division of the total average time or errors involved in practice. The ultimate purpose is to find the ratios between average individual time per trial and the total average time per trial. Since improvement is very slow as subjects approach their physiological limit, an uncalibrated measuring device would obscure the improvement which still takes place.

4. The next step is to compute individual records. First, the raw data for each subject on trials 1-100 are examined for record performances. Every time after the first trial that a new record score is made, i.e., one falling in a lower one of the chosen intervals given in Table 3, it is underscored (indicated by bold face in Table 2). It will be seen by referring to the chosen intervals in Table 2, p. 263, that new record scores falling into consecutive intervals occur in the following trials for Subject I: 2, 3, 7, 12, 17, 75, etc. Individual cumulative totals are next obtained by adding down to (but not including) a new record score. This record score, however, is included in the next total so that all scores enter into the computation. Thus for Subject I the totals obtained from Table 2 are 105, 103, 161, 121, 90, 828.6, etc.

5. The individual records are now converted into group averages. Group totals are determined for each chosen interval (Table 2) by adding up individual totals falling within such an interval. The resulting data represent the total amount of time or work required by the group to progress from one interval on the scale to another.

6. Next divide the group totals by the number of seconds or errors in the step interval, and in turn divide this quotient by the number of subjects at work in that interval (see Table 3). This gives the average individual number of seconds necessary to eliminate one second of the time per trial, or, stated in other words, it gives the average individual number of seconds necessary to progress by one second on the time scale. It is also possible to find the number of trials necessary for eliminating a second or an error by the same technique described above. Data so derived give us the inverse of the conventional time-per-trial or error-per-trial computation.

We have outlined a technique which treats units of measurement in terms of like function, i.e., time per time or accuracy per accu-

TABLE 3

AVERAGE INDIVIDUAL NUMBER OF TIME OR ERROR UNITS PER TRIAL TO PROGRESS ONE TIME OR ERROR UNIT PER TRIAL ON THE TOTAL PERFORMANCE SCALE

Time				Errors				Time in secs. plus errors			
Interval in secs.	Subjects at work	Av. sec. per sec.	Av. trial per sec.	Interval in secs.	Subjects at work	Av. error per error	Av. trial per error	Interval in secs.	Subjects at work	Av. unit per unit	Av. trial per unit
245-145	11	1.99	.010	200-100	12	1.78	.013	409-209	11	1.90	.0059
145-95	17	2.87	.028	100-50	16	1.99	.028	209-109	24	3.73	.0291
95-45	40	3.23	.057	50-25	40	4.75	.205	109-59	40	5.20	.069
45-40	40	5.66	.12	25-20	40	7.77	.485	59-49	40	8.06	.150
40-35	40	8.32	.19	20-15	40	20.88	.944	49-39	40	15.80	.328
35-30	40	14.12	.33	15-10	40	40.17	1.90	39-29	40	28.70	.945
30-25	40	18.50	.59	10-9	40	65.85	2.45	29-19	40	69.85	1.283
25-20	40	25.49	.99	9-8	36	69.20	3.72	19-17	32	225.50	10.40
20-15	40	38.19	1.87	8-7	34	72.85	4.76	17-15	21	277.10	11.41
15-10	40	70.10	4.68	7-6	31	79.30	5.00	15-13	16	286.30	13.08
10-9	38	153.00	11.00	6-5	28	81.40	6.82	13-11	7	295.70	16.80
9-8	23	170.20	14.50	5-4	25	88.01	7.60	11-9	4	304.40	18.82
8-7	20	186.30	18.10	4-3	23	91.20	8.00				
7-6	18	213.00	23.00	3-2	22	95.08	14.41				
6-5	11	230.00	26.50	2-1	20	106.50	17.27				
5-4	4	260.00	29.20	1-0	10	228.10	30.15				

racy. It establishes the difficulty at the various stages of practice since it computes the actual amount of time and the number of errors necessary to progress from one achievement interval to another. Finally, it bases improvement upon overlapping achievement scores and thus gives a common zero for individual decelerations of improvement.

Table 3 summarizes the results for mirror-tracing data thus computed. The intervals represent the achievement range in the entire practice series. The interval column for time shows that the poorest initial score is 245 secs. and the best is 4 secs. Errors range from 200-0 and time plus errors from 409-9. In each case there is a wide range of progress as a result of practice. The table shows that in the lowest and highest intervals not all of the subjects are included. This means that there were subjects whose initial time and error scores were better than 245-95 and 200-100 respectively. There were also subjects whose final time scores were poorer (greater) than 10-4 and whose error scores were poorer than 9-0. An examination of the table would make it appear that the range of overlapping achievement, where all 40 subjects are involved, is comparatively small. However, since initial learning is so rapid in mirror-tracing, it took the poorest subject only a few trials to reach the initial achievement of the best subject. All subjects were represented by time scores on trials 5-55, by error scores on trials 4-71, and by time-plus-error scores on trials 5-63. In referring to Figure 5, where achievement is pictured in terms of time plus errors per trial, it will be seen that a good portion of the progress for the group takes place in these trials.

Column 3 in Table 3 gives the average individual number of seconds required to eliminate one second per trial. Let us suppose that these figures represent increments on the slope of a hill which the subjects under consideration are climbing. At the foot of the hill (initial practice period) the incline is small but it increases in steepness so that the various subjects, according to their individual physiological limits, reach a point where they cannot proceed any farther. Thus considered, we could translate the time data in Table 3 to mean that between the 245-145 level the average subject consumes 1.99 secs. to climb one unit interval; at the 145-95 level, he takes 2.87 secs.; until at the 5-4 level he takes 260 secs. Similar conditions prevail for errors. The number of errors

necessary to eliminate an error increases with practice or with the decreasing size of the achievement score. If we compute the average number of trials necessary to eliminate an average second or an error per subject as shown in columns 4 and 8, we again see a great and consistent increase in the amount of work necessary for the reduction of time and errors. It may thus be concluded in accordance with our findings that the *increments in time necessary to achieve a given time vary inversely with the size of the score of time achieved.* Likewise, *increments in accuracy necessary to achieve a given accuracy vary inversely with the size of the accuracy score achieved.* Since the figures in the third columns for time, error, and time-plus-error data represent a relationship of average individual performance per trial to total performance per trial, they are a measure of what occurs between achievement scores. Stated in other terms, these figures represent the average individual amount of time per trial actually consumed and the average individual number of errors per trial actually made while progressing from one achievement score to another. *They may therefore be regarded as expressing the difficulty of achievement scores. We have thus derived an empirical measure for difficulty in learning.* When, for example, time achievement scores range from 245-145 seconds it is comparatively easy for the average subject to eliminate a second (1.99 seconds); however, when time achievement scores range from 5-4 seconds the difficulty has increased to a ratio of 260/1.99.

The time and error averages have been plotted in Figure 3 to present a clearer picture of the increasing difficulty with increasing achievement. On the abscissa is given the total performance range for time and errors. On the ordinate is shown number of seconds or errors per subject required to progress (eliminate) one second or error on the total range. It will be seen that curve I, which represents time, rises gradually at first and then rapidly as achievement approaches the physiological limit. The error curve (curve II) definitely shows a similar trend. As stated before, the averages in the curves (taken from Table 3) represent the actual amount of time and number of errors falling between achievement scores. The curves may, therefore, be considered time and work curves in the true sense. And, since they express the amount of time consumed and the number of excessive movements (errors) made be-

tween achievement scores, they may be called difficulty curves. It is of importance that the appearance of the curves is what one would expect in a difficulty curve for a motor task. Increasing difficulty is correlated with decreasing achievement. In other words, a difficulty curve gives us a positively accelerated or inverted learning curve. This may be regarded as furnishing empirical evidence for its validity.

After an examination of these data it at once becomes clear why we cannot predict from unweighted scores what effect practice will have upon individual differences. Let us suppose that the averages in the curves represent hurdles and that 40 subjects are asked to jump over as many of them as they can in an allotted number of trials. The poorest subject, starting with the lowest hurdle, will be able to clear a large number of them in the course of practice. However, if difficulty is taken into account, he will not accomplish any more than—if indeed as much as—a subject who starts practice with a more difficult hurdle and clears a smaller number of them. It is obviously absurd to stress the number of hurdles a given subject can pass irrespective of their difficulty, since in practical life we place great emphasis upon accomplishing the difficult. An oversight of this missing link, the difficulty factor, will probably prove to be the chief cause for the diversity of opinion among investigators concerning the effect of practice upon individual differences. For example, the negative correlations between initial achievement scores and percentage of gain in achievement, reported in the literature (25, 18, etc.), have been interpreted on face value to mean a decrease in individual differences as a result of practice. Such an interpretation, however, is valid only if limited strictly to differences in achievement and without reference to the ultimate significance of such differences. If, in accordance with the cited interpretation, we take the size of achievement scores which a subject has reached as indicative of his improvement (as has been the custom) then the subject with a poor initial score will obviously progress a greater distance on the total range covered by the group than a subject starting with a higher score. This is true because poor initial achievement means progress over the relatively easy portion of the total range. The result is a negative correlation between initial achievement and percentage of final gain in achievement. If, however, we take the height of the hurdle, or the difficulty of

a given score, into consideration the reported negative relationship no longer holds. Upon considering the weight of achievement scores, it becomes obvious, as seen in Figure 3, that small differences during final practice are equal to greater differences during initial practice.

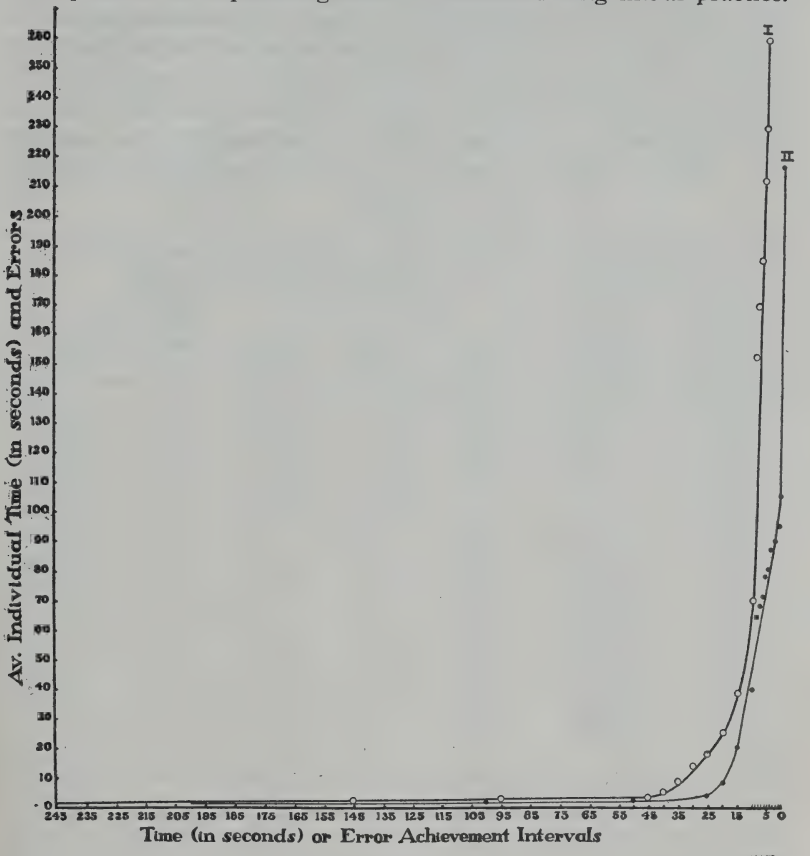


FIGURE 3

THE AVERAGE INDIVIDUAL NUMBER OF SECONDS OR ERRORS PER TRIAL NECESSARY TO PROGRESS (ELIMINATE) A SECOND OR AN ERROR PER TRIAL ON SUCCESSIVE INTERVALS OF THE PERFORMANCE SCALE

On the abscissa is given the total performance range for time (in seconds) and errors per trial. On the ordinate are given the average individual time (in seconds) and errors per trial necessary to progress (eliminate) one second or error per trial. Curve I represents time; Curve II, errors.

We shall next describe how difficulty values are obtained for achievement scores.

Having obtained averages (Table 3) indicating the difficulty at the various levels of achievement, it becomes possible to find difficulty values for raw individual achievement scores at the respective stages of progress. This is done by looking up in Table 3 the levels of difficulty into which the raw scores at the various stages of practice fall. Individual raw-score difficulty values were thus obtained for achievement scores in each fifth trial including the first and the last. Table 4 illustrates for 2 of the 40 subjects how the difficulty

TABLE 4

Taken from columns 1 and 3, Table 3		Raw achievement scores and weights for each fifth consecutive trial				
Achievement intervals (secs.)	Difficulty values (secs.)	Trials	Subj. I raw scores (secs.)	Weights (secs.)	Subj. II raw scores (secs.)	Weights (secs.)
245-145	1.99	1	105	2.87	219	1.99
145-95	2.87	5	41	5.66	54	3.23
95-45	3.23	10	22	25.49	31	14.12
45-40	5.66	15	16	38.19	20	38.19
40-35	8.32	20	15	70.10	23	25.49
35-30	14.12	25	14	70.10	18	38.19
30-25	18.50	30	15.5	38.19	15	70.10
25-20	25.49	35	16	38.19	15	70.10
20-15	38.19	40	18	38.19	13.5	70.10
15-10	70.10	45	13.4	70.10	12	70.10
10-9	153.00	50	14.3	70.10	14	70.10
9-8	170.20	55	13.5	70.10	12	70.10
8-7	186.30	60	13	70.10	12.5	70.10
7-6	213.00	65	11	70.10	9.5	153.00
6-5	230.00	70	12	70.10	13	70.10
5-4	260.00	75	10	153.00	12	70.10
		80	9.5	153.00	11	70.10
		85	9.5	153.00	11	70.10
		90	8	186.30	10	153.00
		95	9	170.20	10	153.00
		100	9	170.20	9.2	153.00

values were obtained. In columns 1 and 2 (taken from Table 3) appear the achievement intervals and their difficulty values. As stated before, these difficulty values describe the average time units per individual per trial necessary to progress from one time interval to another. In columns 4 and 6 appear the raw scores for two subjects taken from each fifth consecutive trial of their total per-

formance. The fifth and last columns give the difficulty values of these scores. They have been read off directly from columns 1 and 2. For example, Subject I, made a time score of 105 secs. in the first trial, consequently the difficulty value is 2.87 secs., as this is the average individual time required to progress from the 145-second to the 95-second achievement level. Table 4 is given here merely to show how difficulty scores were obtained. Individual difficulty scores were thus read off from raw scores (for consecutive fifth trials) for the whole group. This resulted in a distribution for each of the 21 trials under consideration. Averages and S.D.'s were then computed for the difficulty values in each trial.[2]

The resulting data for time are presented in Figure 4. The abscissae give the trials and the ordinates, time in seconds. Curves I and II represent average achievement scores and their S.D.'s per trial, respectively, for unweighted data. They are presented here for purposes of comparison. Curve III gives averages for the difficulty values. In other words, it expresses the difficulty weights of respective averages in curve I. Its significance lies in the fact that it gives the average number of time units (seconds) actually consumed between the T averages in consecutive trials (see the discussion on the meaning of difficulty values). Difficulty time scores will henceforth be referred to as DT. Achievement time scores will be designated by the symbol T. It will be seen that the DT averages in Figure 4 go up at an essentially uniform rate, whereas Figure 3 shows that the difficulty curves rise very abruptly at the end. It must, however, be remembered that Figure 3 gives the difficulty curves for each interval of achievement in terms of the average individual time and errors necessary to reach that interval. Figure 4, on the other hand, represents the average individual difficulty of the actual time achievement scores made by the 40 subjects on each fifth trial from 1-100. Hence the data in Figure 3 represent average individual difficulty norms based upon achievement intervals, whereas the data in Figure 4 represent average individual difficulty values for raw achievement scores based upon trials. Having estab-

[2]Coefficients of variation were also computed but are not presented here since S.D.'s/total av. give figures which run parallel with the S.D.'s (see Figure 2). On the other hand, S.D.'s/av. per *trial* are not considered valid. When this latter formula is applied to the difficulty-of-achievement scores the coefficients fall within a straight line in the range of overlapping achievement, thus indicating constancy of individual differences with practice.

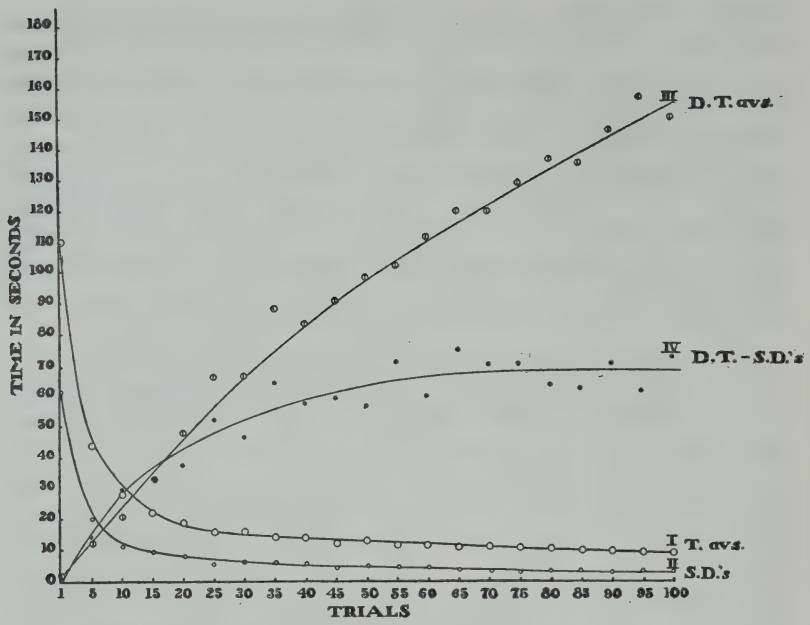


FIGURE 4

AVERAGE ACHIEVEMENT SCORES FOR TIME AND THEIR DIFFICULTY VALUES

On the abscissa are represented successive fifth trials; on the ordinates, time in seconds. Curve I gives achievement scores; and curve II, their S.D.'s. Curve III shows the average difficulty values of raw achievement scores; and curve IV, their S.D.'s.

lished difficulty values for achievement scores, it becomes necessary to consider them in relation to individual achievement from trial to trial in order to deal with the problem of individual differences. Consequently, the absence of a sharp inflection in the curves of Figure 4 results in part from the fact that the raw achievement curve (curve I) is gradual in character and shows no sudden break and also to the smoothing effects of averaging scores. Figure 4 shows that, whereas both the T and DT averages are negatively accelerated, the DT averages increase while the T averages decrease in size with increase in number of trials. As previously stated, this means that difficulty is inversely related to the size of the achievement scores. *Consequently the smaller the achievement scores the more difficult is the achievement attained.* An oversight of this

factor has resulted in much confusion in interpreting the data of studies on individual differences. This finding obviously explains why time-per-work and work-per-time computations give different results, since the one leads to an increase and the other to a decrease in the size of achievement scores, and consequently in the size of their differences.

The next question is how variable are the DT scores? Furthermore, do the individual differences in actual amount of time required to progress from one T score to another increase or decrease with practice? The answer to these questions may be obtained from curve IV which gives the S.D.'s for DT scores. It will be seen that the differences increase with practice up to about the 55th trial and remain constant thereafter. A comparison of T scores and their S.D.'s and DT scores and their S.D.'s shows that the former

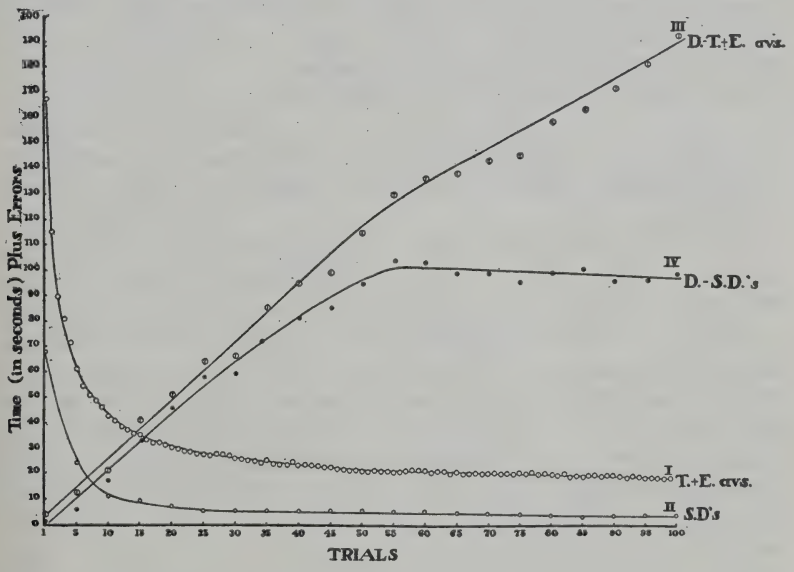


FIGURE 5

AVERAGE ACHIEVEMENT SCORES IN TIME PLUS ERRORS AND THEIR DIFFICULTY VALUES

On the abscissa are represented successive fifth trials; on the ordinate, time (in seconds) plus errors. Curve I gives the achievement averages for each trial, and curve II gives the S.D.'s in successive fifth trials. Curve III gives the average difficulty values of raw achievement scores, and curve IV gives their S.D.'s.

converge and the latter diverge. This is undoubtedly due to the previously reported inverse relationship between the size of T and DT scores, i.e., as T scores decrease their differences also decrease, and as DT scores increase their differences likewise increase. The result is that averages and S.D.'s converge in one case (T scores) and diverge in the other (DT scores).

Figure 5 should, with a few exceptions, be read like Figure 4. One exception is that the data consist of time-plus-error scores and another that T averages are here given for each trial. The T averages (curve I) and their S.D.'s (curve II) yield strikingly smooth curves. This is probably in part due to the type of motivation used. The DT scores in curves III and IV were obtained in the same manner as the DT scores in Figure 4. It will be seen that the results also agree in most essentials with those of Figure 4. Curve III gives the DT averages for each fifth successive trial and curve IV gives their S.D.'s. As in Figure 4, the S.D.'s go up with practice to the 55th trial, thus indicating (to the 55th trial) increased individual differences in difficulty scores as a result of practice. However, the final answer to the question as to the effect of practice upon individual differences awaits the actual weighting of achievement scores by difficulty, a task which will now be undertaken.

Progress in achievement and difficulty have been shown to be inversely related. This would naturally suggest a linear relationship between difficulty and achievement, providing their ratios are constant. If it could be shown that the inverse relation between size of achievement score and difficulty is a linear function or that weighted scores progress at constant ratios from one point on the performance scale to another, we would have an indication for the trend of improvement. Having determined the trend for improvement, it would be more nearly possible to establish a trend for variability, since without reasonable evidence for a trend in improvement we naturally remain in total darkness as to the trend of variability. We shall now present data for weighted achievement scores. Results for the time data are presented in Figure 6. In this case the best achievement scores in each five successive trials, instead of each fifth successive trial as (see p. 270) for the data in Figure 4, constitute the raw data.[3] Curve II in Figure 6 gives the DT

[3]Out of each 5 trials select the best record for each subject. The average of these scores represents T in the figure. Get the DT for each individual record and their average. This represents DT in the figure.

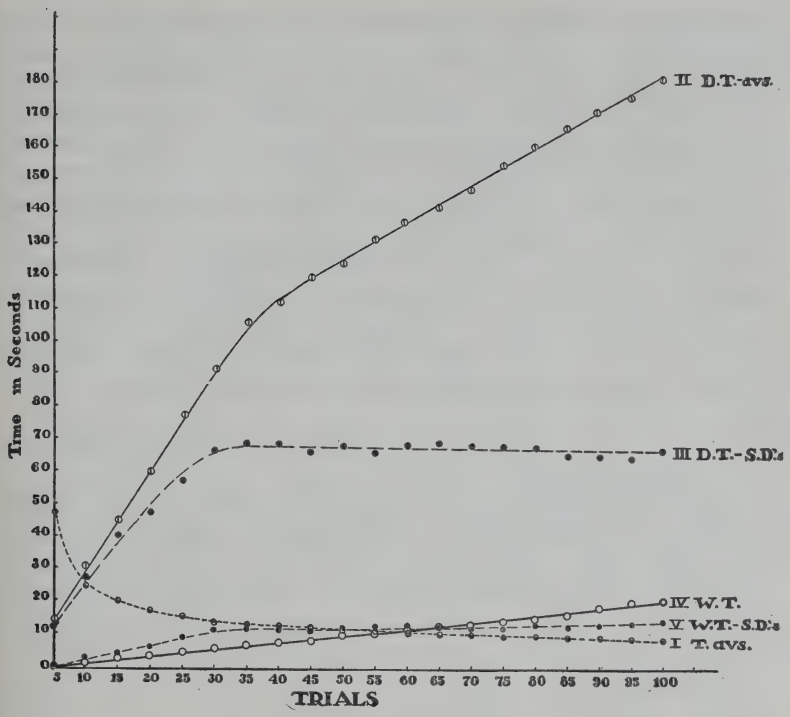


FIGURE 6

WEIGHTED ACHIEVEMENT SCORES AND THEIR S.D.'S AND THE DATA FROM WHICH THEY WERE DERIVED

On the abscissa are represented trials, on the ordinate, time in seconds. Curve I gives the best average achievement scores in successive fifth trials. Curve II gives the difficulty values of the achievement scores in curve I, and curve III gives the S.D.'s for the difficulty values. Curve IV gives average achievement scores weighted for difficulty. It was obtained by dividing individual difficulty scores by individual achievement scores in respective trials and then finding group averages. Curve V gives the S.D.'s for weighted achievement scores.

averages and curve III their S.D.'s. These scores were derived in the same manner as the DT scores presented in Figures 4 and 5, and, as will be seen, their trend is similar. Curve I gives the achievement or T averages. Curve IV represents ratios of difficulty in time and achievement in time. It was found by dividing the DT averages in curve II by respective T averages in curve I.

This gives us a ratio of weighted time and achievement-time or, in other words, *weighted achievement scores.*[4] They will be designated by the symbol WT. The resulting weighted achievement scores tend to fall in a straight line. Such a linear relationship may be taken as further proof for the validity of the obtained difficulty values. Logically, an inversely corresponding negative and positive trend of acceleration in the size of achievement and difficulty scores respectively should give a linear ratio. Curve IV lends support to such a relationship. It further establishes a formula for the trend of improvement on the performance scale, namely, difficulty values in successive trials divided by achievement scores in respective trials equals rate of progress or a trend for improvement. Applied to our time data, the expression reads: $\frac{DT \text{ of successive trials}}{T \text{ of respective trials}}$ = rate of progress from trial to trial or a trend for improvement in the elimination of time per trial, where DT stands for difficulty of achievement time and T stands for achievement time. Since the inverse relation between size of achievement scores and difficulty scores is a linear function, as shown in Figure 4, we are naturally led to the conclusion that the trend for improvement in mirror-drawing is linear or that it progresses at equal ratios with practice.

The establishment of linearity in progress during practice is of great importance to all learning problems since it indicates uniformity for successive increments of improvement. Stated in other words, it shows that progress is a constant ratio of difficulty and achievement for equal amounts of practice. This finding, although quite foreign to previous reports, is in agreement with the linear relationship which has been found to exist between chronological and mental growth (35, 36). Apparently, what has been established for a long period of time holds also for shorter practice periods, namely, linear progress in ability. It should of course be kept in mind that the term "linear," as here used, does not mean that all subjects progressed at an equal speed. It means that their decrement (reduction of time and errors) ratios of improvement remain

[4]The data represented in Figure 4, curves I and III, and Figure 5, curves I and III, have been similarly weighted and the results (not presented here) verify in all essentials those of curve IV in Figure 6.

equally decelerated. Such a finding, if generally established for other activities, would of course modify many accepted concepts concerning the effect of practice upon improvement, e.g., rapid initial improvement, the meaning of individual plateaus, etc. It should further be pointed out that, in basing our findings of the learning process upon scores which have been weighted on the basis of empirically derived difficulty units, we obtain license for the use of the term "improvement" in connection with group performance. As pointed out on page 258, improvement is a change or trend in achievement. This trend, which constitutes improvement, has been measured by weighting achievement in terms of empirically derived difficulty scores as described in this paper. Let us suppose that the subjects engaged in mirror-tracing represent motor cars traveling at different speeds. One starts at 15 miles per hour, another at 30, another at 45, etc. As a group, these cars travel over a distance of 100 miles but they start and stop at different points on the scale. The 15-miles-per-hour car starts on the 0-mile mark, the 30 on the 5, the 45 on the 10, etc. They come to a standstill at the following distances of the scale: the 15-miles-per-hour car at 90, the 30 at 95, the 45 at 100. Each car decelerates at a constant rate, but different cars decelerate at different rates. The ratios of the distances of the moving cars from their starting-points will decrease with time and per unit of distance traveled. However, the average distances between individual cars will remain a constant ratio from second to second and from unit to unit of distance traveled. If, for this hypothetical problem, we weight distance intervals for individual rates of deceleration, we obtain a linear relationship like that obtained for mirror-tracing.

But what bearing does linearity in improvement have upon the problem of individual differences? It is to be considered as furnishing a trend for improvement upon which we may base variability. In order to solve the effect of practice on individual differences we need to know what general trend performance will take in the light of average difficulty as the group approaches its physiological limit. Once this trend has been established, measures of variability may be applied to indicate individual differences.

S.D.'s have been computed for WT data and are presented in curve V, Figure 6. They were obtained as follows: The individual DT scores for each successive five trials were divided by the

individual T scores for each respective five trials. This results in 20 distributions, each representing group performance in best scores for successive five trials. For these S.D.'s were computed. Coefficients of variability are not presented for these data since, as pointed out below, the formula S.D. for each trial/total av. results in scores which run parallel with their respective raw S.D.'s. It will be seen that the S.D.'s in curve V increase in the initial trials and then remain practically constant. A comparison of the 1st and 100th trials shows that the S.D.'s have increased from about .5 to 14. The importance of the increase in S.D.'s is enhanced by the fact that it follows in general the increase of the weighted score averages. In accordance with these findings we may conclude that *when units of measurement in mirror-tracing are weighted for difficulty individual differences tend to increase with practice.* It should further be noted that the S.D.'s fall in a comparatively straight line beyond the initial increase. This suggests a linear increase for individual differences during practice. Hence equal amounts of practice on a performance scale tend to result in a linear trend in the ratios of differences. Or, stated in the reverse, equal amounts of improvement tend to be equally difficult for all of the subjects in the group. However, a linear trend for individual differences with practice would presuppose a perfect correlation between initial achievement and gross gain in achievement, i.e., each subject would retain his relative achievement rank no matter how small the difference. The previously reported correlation of .98 (p. 262) between initial achievement and gross gain would practically fulfill this requirement. No measure of human traits could, of course, be expected to be perfectly reliable. The lower correlations previously reported in the literature (18, 25, etc.) are probably due to less uniformity in motivation. They may also be due to many other factors, such as unequal initial orientation, unequal loss of interest, unequal degrees of fatigue, etc.

The findings of this paper are not in agreement with Reed's (24, 25) statement that correlations between percentage of improvement and initial performance are a better measure of individual differences than correlations between gross gain and initial performance. The latter gives us at least a measure of achievement. But percentage of gain, or any measure of relative gain, takes for granted equality of unweighted units of measurement. I have shown that

the problem of improvement and individual differences in improvement is inextricably linked up with the difficulty factor. I have further shown that difficulty is inversely related to the size of the achievement score. My findings are in closer agreement with those of Thorndike and his students. In his early paper of 1908 Thorndike (29) says:

> It seems extremely probable from many facts of dynamic psychology that the man who has the capacity to improve to a given small degree more quickly than another, should also improve more quickly to the next degree and should also by the by be capable to improve to a given higher degree if given the maximum of efficient training. (p. 384)

In a later paper Thorndike (32) says:

> The effect of equalizing opportunity is thus to increase individual differences. This result now found with many different functions, furnishes perhaps the strongest argument that differences in achievement are largely due to differences in original capacity. (p. 553)

The first part of the latter quotation is substantiated by my data. However, the last part of the statement cannot be tested by any data of this sort since, unless the subjects are carefully chosen for the sole purpose of answering the question, we can never know whether improvement during practice is due to nurture or to nature.

I believe to have accomplished the objectives set for this investigation. Scores have been weighted for difficulty and probable determination or zeros have been obtained by working with achievement scores of overlapping difficulty. In addition, units of measurement have been treated in terms of ratios of like function, thus simplifying interpretations of the results.

It should be pointed out that Thurstone's (34, 35) method for finding the absolute zero in intelligence measurement accomplishes in general the same objectives in the treatment of group data as the method described in this paper. But, as Thurstone (34, 35) himself has pointed out, it is valid only when based upon absolute scaling. Absolute scaling in turn assumes normal distributions in achievement for every trait or group of traits appearing on the entire scale. Consequently, since, during learning, separate trials in achievement do not ordinarily yield normal distributions, Thurstone's method is not adapted for the computation of learning data.

IV. Summary and Conclusions

The problem concerning the effect of practice upon individual differences has been investigated in the light of a statistical technique which takes into account the following requirements for the solution of the problem: (1) Units of measurement are treated in terms of like ratios, such as time per time or work per work; (2) achievement scores have been weighted for difficulty; (3) units of measurement have been based upon a scale on which scores have a common zero.

Data were gathered in a preliminary experiment in a digit-substitution test for a reexamination of accepted methods of computation, and, if possible, to show why the various methods reported in the literature give different results. It is shown that the conventional method of computing units of measurement in terms of unlike function is inadequate for detecting what occurs between achievement scores and consequently cannot be used for measuring improvement. Hence we cannot regard the negative correlation between initial achievement or intelligence and the percentage of gross gain in achievement, which was found in the preliminary experiment and which has been reported in the literature, as finally established. In order to determine relative variability of improvement with practice it is necessary to compute the difficulty values of successive achievement scores and to base comparison upon common levels of performance.

A second experiment in mirror-tracing was conducted under intense motivation with these objectives in view and the conditions under which they were accomplished have been described. The steps necessary for the derivation of equal ratios in units of measurement have been outlined. Difficulty measures were obtained by finding the average individual time or work units per trial necessary to progress one such average unit on the total time or work scale. Individual increments of progress are thus always taken as ratios on the total performance scale. The average time or work unit necessary to progress from one achievement score (which overlaps in size for all subjects) to another is taken as the difficulty measure. This furnishes an empirical measure for difficulty in learning.

Analysis of the results leads to the following conclusions:

1. A comparison of achievement averages with difficulty averages

(time or work units lying between achievement scores) shows that the two are inversely related. If pictured in terms of learning curves, achievement averages give the typical negatively accelerated curve, whereas the difficulty averages give a positively accelerated or inverted learning curve. This may be taken as evidence for the validity of our difficulty measures. Further comparison of difficulty and achievement scores shows that the *increments in time or work necessary to achieve a given time or work score vary inversely with the size of the score achieved.*

2. *Smaller achievement differences during final practice are equal to greater differences during initial practice when the difficulty factor is taken into consideration.* Consequently, the decrease in achievement scores and their individual differences, which has been reported in a preponderance of recent studies in this field, creates the illusion that individual differences decrease with practice. It also creates the illusion that the negative correlations between initial achievement and percentage of gross gain in achievement in a motor task should be taken to mean that individual differences decrease with practice. The illusion results in either case from ignoring the difficulty factor.

3. When achievement scores are weighted for difficulty, the group averages, considered trial by trial, tend to fall in a straight line. Such a linearity may be taken as further proof for the validity of the obtained difficulty values. *Logically, an inversely corresponding negative (achievement averages) and positive (difficulty averages) acceleration should give a linear ratio.*

4. Constancy in improvement with practice is in agreement with the linear relationship which has been found to exist between chronological and mental growth (35, 36).

5. The empirical relationship between achievement scores and difficulty scores furnishes a formula for the trend of improvement, namely, *difficulty in terms of actual time or work units in successive trials/time or work achieved in respective trials = rate of progress on the scale or a trend of improvement. Improvement may thus be defined as a trend from worse to better or more to less in achievement which may require much or little work.*

6. The inverse relation between size of achievement score and difficulty is a linear function.

7. *Equal amounts of practice on a performance scale results in*

a linear trend of individual differences. Stated in the reverse, equal amounts of improvement are equally difficult for all the subjects in the group.

8. *When units of measurement in mirror-tracing are weighted for difficulty individual differences tend to increase with practice.*

References

1. Argelander, A. Zur Frage der Übungsfähigkeit. *Psychotechn. Zsch.,* 1928, **3**, 141-147.
2. Barlow, M. C. Individual differences as affected by continued practice. M.A. thesis, George Peabody Coll. Teach., 1924.
3. Chapman, J. C. Individual differences in ability and improvement and their correlations. *Teach. Coll. Contrib. Educ.,* 1914, No. 63. Pp. 45.
4. ————. Learning curve in typewriting. *J. Appl. Psychol.,* 1919, **3**, 252-268.
5. ————. Statistical considerations in interpreting the effect of training on individual differences. *Psychol. Rev.,* 1925, **32**, 224-234.
6. Donovan, M. E., & Thorndike, E. L. Improvement in a practice experiment under school conditions. *Amer. J. Psychol.,* 1913, **24**, 426-428.
7. Ehinger, G. Recherches sur la développement de l'habileté manuelle par la partique d'un metier manuelle. *Arch. de psychol.,* 1927, **80**, 299-319.
8. Eschowitz, N. Zur Frage der Bewährung von Handgeschichlichkeitsproben. *Psychotechn. Zsch.,* 1927, **2**, 23-26.
9. Ewert, P. H. An historical and experimental study of the transfer of training. M.A. thesis, Univ. Kan., 1925.
10. Franzen, R. 'Statistical issues.' *J. Educ. Psychol.,* 1924, **15**, 367-382.
11. Gates, G. S. Individual differences as affected by practice. *Arch. Psychol.,* 1922, **8**, No. 58. Pp. 74.
12. Gemelli, A. Recherches sur le diagnostic de l'habileté motrice. *Rev. de la sci. du travail,* 1929, **1**, 181-196.
13. Hahn, H. H., & Thorndike, E. L. Some results of practice in addition under school conditions. *J. Educ. Psychol.,* 1914, **5**, 65-84.
14. Hildebrandt, H. Übung und Eignungsprüfung. *Indus. Psychotechn.,* 1926, **3**, 96-109.
15. Hunter, W. S. Habit interference in the white rat and in human subjects. *J. Comp. Psychol.,* 1922, **2**, 29-56.
16. ————. Learning: II. Experimental studies of learning. In *The foundations of experimental psychology,* ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1929. Pp. 564-627.
17. Kellner, H. Neun Jahre Prüferfahrungen in der Berliner Metallindustrie. *Indus. Psychotechn.,* 1928, **5**, 46-47.

18. KINCAID, M. A study in individual differences in learning. *Psychol. Rev.*, 1925, **33**, 34-53.
19. MATTSON, M. L. The relation between complexity of the habit to be acquired and the form of the learning curve in young children. *Genet. Psychol. Monog.*, 1933, **13**, 299-398.
20. PETERSON, J. Experiments in ball tossing: the significance of learning curves. *J. Exper. Psychol.*, 1917, **2**, 178-224.
21. ———. Thurstone's measures of variability in learning. *Psychol. Bull.*, 1918, **15**, 452-455.
22. PETERSON, J., & BARLOW, M. C. The effect of practice on individual differences. *27th Yrbk. Nat. Soc. Stud. Educ.*, 1928, Pt. 2, 211-230.
23. RACE, H. Improvability: its inter-correlations and its relation to initial ability. *Teach. Coll. Contrib. Educ.*, 1922, No. 124. Pp. 68.
24. REED, H. B. The effect of training on individual differences. *J. Exper. Psychol.*, 1924, **7**, 186-201.
25. ———. The influence of training on changes in variability in achievement. *Psychol. Monog.*, 1931, **41**, No. 185. Pp. xli+59.
26. RUCH, G. M. The influence of the factor of intelligence on the form of the learning curve. *Psychol. Monog.*, 1908, **34**, No. 160. Pp. 62.
27. SNODDY, G. S. An experimental analysis of a case of trial and error learning in the human subject. *Psychol. Monog.*, 1920, **38**, No. 124. Pp. 78.
28. STODDARD, G. D. The problem of individual differences in learning. *Psychol. Rev.*, 1928, **35**, 479-485.
29. THORNDIKE, E. L. The effect of practice in case of a purely intellectual function. *Amer. J. Psychol.*, 1908, **19**, 374-384.
30. ———. Practice in case of addition. *Amer. J. Psychol.*, 1910, **21**, 483-486.
31. ———. Relation between initial ability and improvement in a substitution test. *School & Soc.*, 1915, **1**, 429-430.
32. ———. Notes on practice, improvability and curve of work. *Amer. J. Psychol.*, 1916, **27**, 550-565.
33. THURSTONE, L. L. Variability in learning. *Psychol. Bull.*, 1918, **15**, 210-212.
34. ———. A method of scaling psychological and educational tests. *J. Educ. Psychol.*, 1925, **16**, 433-451.
35. ———. The absolute zero in intelligence measurement. *Psychol. Rev.*, 1928, **19**, 441-453.
36. THURSTONE, L. L., & ACKERSON, L. The mental growth curve for the Binet tests. *J. Educ. Psychol.*, 1929, **20**, 569-583.
37. TINKER, M. A., & FOSTER, W. S. Experiments in psychology. (Rev. ed.) New York: Holt, 1929. Pp. xv+392.
38. VITELES, M. S. Influence of training on motor test performance. *J. Exper. Psychol.*, 1933, **16**, 556-564.
39. WELLS, F. L. The relation of practice to individual differences. *Amer. J. Psychol.*, 1912, **23**, 75-88.

University of Vermont
Burlington, Vermont

L'EFFET DE L'EXERCICE SUR LES DIFFÉRENCES INDIVIDUELLES ÉTUDIÉES AVEC LES MESURES APPRÉCIÉES SELON LA DIFFICULTÉ

(Résumé)

Le problème de l'effet de l'exercice sur les différences individuelles à été étudié au moyen d'une technique statistique qui tient compte des conditions suivantes de la résolution du problème: (1) On traite les unités de mesure en termes des mêmes proportions, tels que temps pour temps ou travail pour travail; (2) on a apprécié les résultats du rendement selon la difficulté; (3) on a basé les unités de mesure sur une échelle sur laquelle les résultats ont le même zéro. Quelques-unes des conclusions les plus frappantes sont les suivantes:

1. Une comparaison des moyennes du rendement et celles de la difficulté montrent que les deux ont une relation inverse. Une analyse plus étendue montre que les accroissements du temps ou du travail nécessaires pour achever un résultat donné de temps ou de travail varient inversement avec la grandeur du résultat achevé.
2. Les plus petites différences du rendement pendant l'exercice final sont égales aux plus grandes différences pendant l'exercice initial quand l'on tient compte du facteur de difficulté. Par conséquent, le décroissement si souvent rapporté des résultats du rendement et leur variabilité crée l'illusion que les différences individuelles décroissent avec l'exercice.
3. Quand on apprécie les résultats du rendement selon la difficulté, les moyennes collectives, considérées épreuve par épreuve, tendent à tomber dans une ligne droite.
4. La relation empirique entre la difficulté et le rendement donne la formule suivante pour l'amélioration: la difficulté dans les unités elle-mêmes de temps ou de travail pour les épreuves successives/le temps ou le travail achevés dans les épreuves respectives est égale à une tendance à l'amélioration.
5. Les mêmes quantités de l'exercice sur une échelle d'exécution montre une tendance linéaire des différences individuelles.
6. Quand on apprécie les unités de mesure dans l'écriture au miroir selon la difficulté, les différences individuelles tendent à s'accroître avec l'exercice.

EWERT

DIE EINWIRKUNG DER EINÜBUNG AUF INDIVIDUELLE UNTERSCHIEDE UNTERSUCHT MITTELS MESSUNGEN DIE IN BEZUG AUF SCHWIERIGKEIT AUSGEGLICHENEN WORDEN SIND

(Referat)

Das Problem der Einwirkung der Einübung auf individuelle Unterschiede ist untersucht worden mit Hilfe einer statistischen Technik, die die folgenden Erfordernisse zur Lösung des Problems in Betracht zieht: (1) Messungseinheiten werden behandelt in Form von Proportionsgleichungen, solche wie Zeit: Zeit oder Leistung: Leistung; (2) Leistungssummen sind in Bezug auf Schwierigkeit ausgeglichen worden [achievement scores have been weighted for difficulty]; (3) Messungseinheiten sind auf einer Skala aufgebaut worden auf der die einzelnen Summen eine gemeinsame Null

haben. Es folgen einige besonders aufschlussreiche Feststellungen:

1. Ein Vergleich des Leistungsdurchschnitte mit den Schwierigkeitsdurchschnitten zeigt, dass die beiden umgekehrt proportional sind. Eine weitere Analyse zeigt, dass die Zunahme von Zeit oder Arbeit, dit notwendig ist, eine gegebene Zeit- oder Arbeitsumme zu leisten, sich in umgekehrtem Verhältnis zum Umfange der geleisteten Summe verändert.

2. Kleinere Leistungsunterschiede am Ende des Messungsverfahrens sind gleich grösseren Unterschieden am Anfang des Verfahrens wenn der Schwierigkeitsfaktor in Betracht gezogen wird. Daher entsteht durch die so oft berichtete Abnahme der Leistungssummen und ihre Veränderlichkeit die Täuschung, dass individuelle Unterschiede in Lauf der Einübung abnehmen.

3. Wenn Leistungssummen in Bezug auf Schwierigkeit ausgeglichen werden, zeigen die Gruppendurchschnitte, von Versuch zu Versuch, betrachtet die Tendenz, in eine gerade Linie zu fallen.

4. Die empirische Beziehung zwischen Schwierigkeit und Leistung liefert folgende Formel in Bezug auf Besserung: Schwierigkeit in realer Zeit oder Arbeitseinheit für aufeinander folgende Versuche in ihrem Verhältnis zu Zeit oder geleisteter Arbeit in entsprechenden Versuchen ist gleich einer Neigung zur Besserung.

5. Der gleiche Betrag an Einübung, auf einer Leistungsskale gemessen, ergibt eine lineare Richtung individueller Unterschiede.

6. Wenn die Messungseinheiten beim Anfertigen von Spiegelabzeichnungen in Bezug auf Schwierigkeit ausgeglichen werden, zeigen individuelle Unterschiede die Tendenz, im Lauf der Einübung zuzunehmen.

EWERT

THE ALGEBRA OF LUMINOSITY IN RELATION TO WHITE COMPLEMENTARY PAIRS OF WAVE-LENGTHS*

From the Department of Psychology of The Ohio State University

R. D. Williams

The following symbols used in this discussion are defined thus:
L_t = total luminosity.

L_W = the whiteness component of the experience generated by any given wave-length.

L_H = the resulting hue experience which is generated by any given wave-length.

The luminosity resulting from any λ is a function of the sum of the ordinates of the three excitation curves.

$$L_{(\lambda)} = V_{(\lambda)} + G_{(\lambda)} + R_{(\lambda)}.$$

One of the properties of the psychological visual experience is luminosity for the particular λ. There is the luminosity of brilliance of a particular hue and there is also the luminosity of the white. Degree of brilliance of both are universal forms of experience; a red may be more or less bright than another, and a gray may be more or less bright than another gray. Since the visual experience embodies both forms of luminosity, and since light waves are the only normal stimulus to visual experience, it follows that light always produces at least two effects. These are the experience of hue and the experience of white. There is luminosity of hue and luminosity of white. These are given by the following functions.

$$L_H(\lambda) = \begin{matrix} L_{H_1}(\lambda) = \left(V_{(\lambda)} - R_{(\lambda)}\right) + \left(G_{(\lambda)} - R_{(\lambda)}\right) \\ L_{H_2}(\lambda) = \left(R_{(\lambda)} - V_{(\lambda)}\right) + \left(G_{(\lambda)} - R_{(\lambda)}\right) \end{matrix}$$

$$L_{white}(\lambda) = \begin{bmatrix} L_{W_1}(\lambda) = 3R_{(\lambda)} \\ L_{W_2}(\lambda) = 3V_{(\lambda)} \end{bmatrix}$$

*Recommended by George F. Arps, H. E. Burtt, and Samuel Renshaw, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, March 14, 1933.

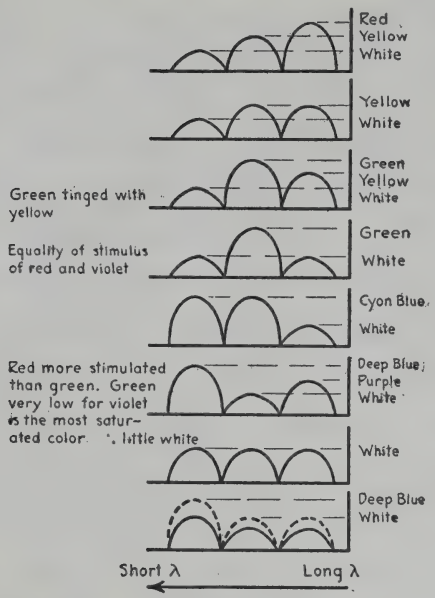


FIGURE 1

The equations defining L_{W} (λ) follow directly from definition, in that white is a measure of the equal excitation of the elements of the three groups. [At this point see Figure 1, Allen's (1) graphic way of depicting this fact.] Let the reader ponder carefully over this way of determining quantitatively the value of the luminosity of white, as this will be central in the interpretation of later concepts.

Luminosity has the character of a vector. In one case luminosity is expressed as a set of triplets $\left(V_{(\lambda)} + G_{(\lambda)} + R_{(\lambda)}\right)$ and in another it is expressed as a doublet. The transformations for these are quite simple and they can be shown to possess all the characters of vector.

Since any specimen of light produces two types of luminosity, the total luminosity is the sum of these two.

$L_t \equiv$ total luminosity

$L_{t(\lambda)} = L_{W(\lambda)} + L_{H(\lambda)}$ (Here luminosity is broken up into two parts.)

$L_{t(\lambda)} = V_{(\lambda)} + G_{(\lambda)} + R_{(\lambda)}$ (Here it is expressed as a triplet.)

In both cases brightness is expressed as a *sum,* whereas hue is expressed as the sum of *two differences.* Whiteness—the *absence of hue*—means that there is an addition of *all* the number of functioning cones. Brightness of a given hue is a measure of the number of receptors that a given λ can set into action. Luminosity, whether it be that of whiteness or that of a hue, is in both cases a function of a sum. Hueness is not dependent on summation but is a matter of the relative amounts of the different cone groups or families that are activated. As each and every λ produces two effects, it is apparent that total luminosity is made up of two parts and it also follows that the luminosity of whiteness is an expression of a balance of the three components, whereas hue is a measure of an imbalance. This latter will be developed below.

The balanced and also the imbalanced character of the effect of cone stimulation can be shown as follows: Suppose we stimulate 50 green cones, 50 blue cones, and 50 red cones. Here the difference between, say, green cones activated and red cones activated is zero; also the difference between the blue cones activated and red cones activated is zero. In these two cases there is not an excess, or imbalance, of any one cone family over the other two.

$L_H(\lambda)$	$L_W(\lambda)$	V	G	R	Balanced amount	Imbalanced amount
0	150	50	50	50	150	0
30	120	60	50	40	120	30

The above table can be continued to show an extended array of balanced and imbalanced amounts of luminosity. The balanced amount can be determined immediately by taking three times the minimum value given under V, G, or R. Luminosity is the amount of balance (three times the minimum ordinate corresponding to any λ).

The basic functions descriptive of luminosity are these:

(1) $L_{(\lambda)} = V_{(\lambda)} + G_{(\lambda)} + R_{(\lambda)}$ (shows total luminosity in terms of the three receptor groups)

(2) $L_H(\lambda) = \begin{cases} L_{H_1}(\lambda) = \left(G_{(\lambda)} - R_{(\lambda)}\right) + \left(V_{(\lambda)} - R_{(\lambda)}\right) \\ L_{H_2}(\lambda) = \left(G_{(\lambda)} - V_{(\lambda)}\right) + \left(R_{(\lambda)} - V_{(\lambda)}\right) \end{cases}$

(3) $L_W(\lambda) = \begin{cases} L_{W_1}(\lambda) = 3R_{(\lambda)} \\ L_{W_2}(\lambda) = 3V_{(\lambda)} \end{cases}$

It can now be shown that the algebraic processes of addition and of subtraction work here, and therefore have application to the psychological processes that are herein depicted.

(4) $L_{\text{total}} = L_{W_1}(\lambda) + L_{H_1}(\lambda) = 3R_{(\lambda)} + \left\{G_{(\lambda)} - R_{(\lambda)}\right\} + \left\{V_{(\lambda)} - R_{(\lambda)}\right\}$

(5) $L_{\text{total}} = L_{W_2}(\lambda) + L_{H_2}(\lambda) = 3V_{(\lambda)} + \left\{G_{(\lambda)} - V_{(\lambda)}\right\} + \left\{R_{(\lambda)} - V_{(\lambda)}\right\}$

From the right hand side of equation (4) it follows that

(6) $L_{\text{total}} = 3R_{(\lambda)} + G_{(\lambda)} - R_{(\lambda)} + V_{(\lambda)} - R_{(\lambda)} = G_\lambda + B_\lambda + R_\lambda = L_{W_1}(\lambda) + L_{H_1}(\lambda)$

and from (5) it follows that

(7) $L_{\text{total}} = 3V_{(\lambda)} + \left\{G_{(\lambda)} - V_{(\lambda)}\right\} + \left\{R_{(\lambda)} - V_{(\lambda)}\right\} = 3V_\lambda + G_\lambda - V_\lambda + R_\lambda - V =$

$$L_{\text{total}} = 3V_\lambda + G_\lambda - V_\lambda + R_\lambda - V_\lambda = G_\lambda + V_\lambda + R_\lambda = L_{W_2}(\lambda) + L_{H_2}(\lambda).$$

Thus from equations (1), (2), and (3) we deduce the corollary

$$L_\lambda = L_{H(\lambda)} + L_{W_\lambda}.$$

In describing briefly the phenomena of complementariness it is well to re-emphasize the fact that any complementary color does two things. First, it produces a sensation or experience of color or hue; and, secondly, it, at the same time, produces white, which is the absence of hue. Hue and the absence of hue are the inevitable

results of the action of any complementary wave-length. In all cases of complementaries, the hues cancel each other. The situation is such that one is led in the interest of simplification of quantitative description to regard the effect of any wave-length as of the character of a complex number. Here, we only suggest that the problem of complementaries can be best described in terms of complex functions. The working-out of this is left for another study.

The experience of complementaries is one of the most common of all our experiences of vision. We would like to have the classical elementary and fundamental curves of Helmholtz, König, and others to tell us something about complementary colors. They will not do so. They stubbornly refuse to tell anything more than has been put into them. Now all that has been put into them is the body of phenomena arising from color-mixture. We cannot, by any device, insert our white complementary colors into the sensation curves.

White is a peculiar experience. It is peculiar from the psychological point of view and it is also peculiar from the physical point of view. From the psychological viewpoint it is the experience of the absence of something—it is the absence of hue. From the physical point of view it is peculiar in that it is the resultant of the simultaneous action of all the primaries or the resultant of certain pairs of wave-lengths. When all the primaries, or, otherwise stated, all the various receptor groups are equally and simultaneously stimulated, the experience is an experience of white; so it is when the eye is stimulated with appropriate pairs of wave-lengths. In both instances there is a cancellation of hues.

One could think of this in terms of the axis of real numbers. And one could think of one hue as a positive real number and the other as a negative real number. This interpretation, however, breaks down and on further study one finds that what we have here is an experimental field which is very much more easily and fruitfully described in terms of complex numbers, with their real and imaginary components. Work that has already been done shows that this method of analysis can be applied successfully to the explanation of a variety of scattered facts and that in terms of this concept they can be unified by being brought into the compass of a single number picture.

There is a twoness that runs throughout all the phenomena of complementaries. There is also a threeness, but the threeness can be reduced to twoness for all complementaries.

The sum, viz., $L_{H_{\lambda_1}} + L_{H_{\lambda_2}}$, of any two wave-lengths which are complementaries is white. L_H means the luminosity of the hue of the particular wave-length indicated here by the subscript of the H; this subscript being λ_1 or λ_2.

$L_{H(\lambda_1)} + L_{H(\lambda_2)}$ = white. Here the hue of λ_1 has cancelled the hue of λ_2. Now since white results from their summation and the hue of each has been cancelled, we can interpret one component as real and the other as imaginary. Suppose we represent the white component as the real and the hue component as the imaginary. It is then easy to express mathematically relationships of two wave-lengths which are complementaries.

$$\text{Let } A = L_W(\lambda_1)$$
$$C = L_W(\lambda_2)$$
$$jB = jL_H(\lambda_1)$$
$$j(-B) = -jL_H(\lambda_2)$$
$$\therefore \quad A + jB = L_W(\lambda_1) + jL_H(\lambda_1)$$
$$C + j(-B) = L_W(\lambda_2) + j\left(-L_H(\lambda_2)\right)$$

It will be noted that the two imaginary parts are equal in absolute value but are opposite in sign. This is so because of the fact that the hue of λ_1 just cancels the hue of λ_2.

An analogy may be given here that is of great help in getting the student clearly to understand this twoness of the effect of a given wave-length in setting up hue experience and experience of whiteness. Imagine a series of three metallic arms of equal length and as nearly alike in all respects as is possible. Let these arms be tightly joined together at one end and also let them extend out from this meeting point making an angle of 120° with each other. Let the center be connected to a vertical support and also have these arms so fastened that they can rotate until they fall into a position of equilibrium. Imagine now that a weight is fastened at the end of each of these arms. If the weights are each equal to each other it is apparent that there is a condition of balance or equilibrium. Corresponding to this state of affairs is the condition which exists

in the retina when the activity of each of the three receptor groups is exactly equal to that of each of the other two. In this case there is only the experience of white light. (See Figure 2.) Here the

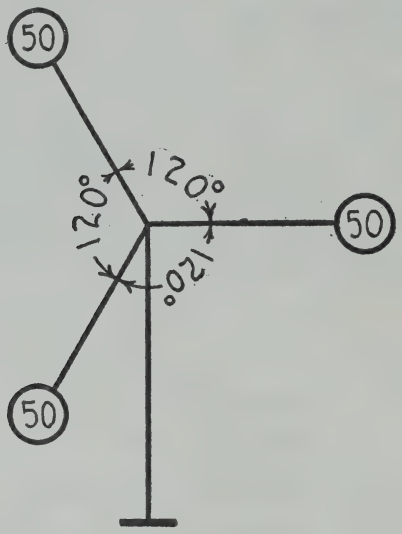


FIGURE 2

arms stay in any position for there is an exact balance and we may interpret luminosity as the sum of the downward pulls of the items given at the end of each arm. In this case there is no imbalance. If we state this in terms of the ordinates of the visibility curves, we would say that all of the ordinates are equal.

In the case where there is an experience of hue the situation is quite different. In this case the ordinates are not equal in magnitude. The experience of hue occurs only when the relative difference of the three ordinates is some positive number. (See Figure 3.) The total sum of the ordinates here is:

$60 + 50 + 40 = 150 = L_t$

$(3)\ (40) = 120 = L_W(\lambda)$ = balanced part

$(60 - 40) + (50 - 40) = 30 = L_H(\lambda)$ = imbalanced character

The imbalanced part of the luminosity is what is effective in set-

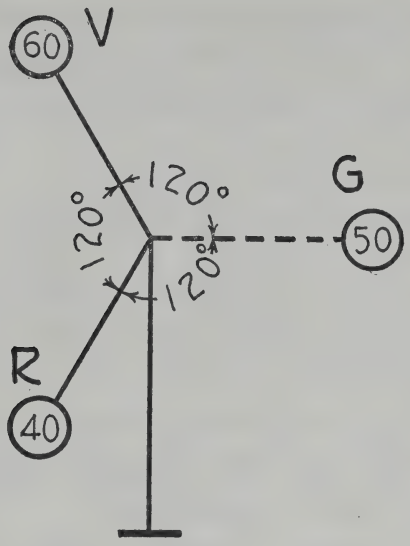


FIGURE 3

ting up the experience of hue. Corresponding to any λ the ordinates of the three primaries express three numbers. The smallest ordinate and an equal part of each of the other two go to make up the balanced part which sets up the experience of white. Equal portions of the three ordinates signify that an equal number of receptors of the three groups is at work.

Luminosity may be expressed in terms of a vector triplet or in terms of a vector doublet in this way:

$$L_{(\lambda)} = V_{(\lambda)} + C_{(\lambda)} + V_{(\lambda)}.$$

When expressed as a triplet, the emphasis is upon the threeness of the receptor groups.

$L_{(\lambda)} = L_W(\lambda) + L_H(\lambda)$ is an expression of the twoness of the effect of retinal stimulation. On the one hand is set up the luminosity of the white, and on the other the luminosity of the hue. So far as brilliance is concerned this is the entire number of effects that the light can generate. Now as we have this twoness exhibited by the luminosity of any wave-length, it is well to isolate this aspect

of the case in our attempt at bringing out just what is involved in the experience of white complementary pairs.

It is perfectly arbitrary what wave-length we start with. Let us indicate this by λ_1. The complementary of this one we shall symbolize by λ_2.

$$L_{(\lambda_1)} = V_{(\lambda_1)} + G_{(\lambda_1)} + R_{(\lambda_1)}$$

Here we are thinking primarily of the receptors:

$$L_{(\lambda_1)} = L_{W(\lambda_1)} + L_{H(\lambda_1)}$$

and this is the theoretical start given us by both Thomas Young and Helmholtz. In this doublet we are thinking more especially of the distinctly psychological aspect of the case.

It follows that

$$L_{(\lambda_1)} = V_{(\lambda_1)} + G_{(\lambda_1)} + R_{(\lambda_1)} = L_{W(\lambda_1)} + L_{H(\lambda_1)}$$
$$L_{(\lambda_2)} = V_{\lambda_2} + G_{\lambda_2} + R_{\lambda_2} = L_{W(\lambda_2)} + L_{H(\lambda_2)}$$

The sum of the above two equations is

$$L_{\lambda_1} + L_{\lambda_2} = L_{W(\lambda_1)} + L_{W(\lambda_2)} + L_{H(\lambda_1)} + L_{H(\lambda_2)}$$

If this sum gives only experience of white it follows immediately that

$$L_{H(\lambda_1)} = L_{H(\lambda_2)}$$

and

$$L_{H(\lambda_1)} + L_{H(\lambda_2)} = 0$$

In terms of vector addition, this means that the imaginary component (hue) of the one is equal, but of opposite sign, to the other.

As given before, let

$$A + jB = L_W(\lambda_1) + jL_H(\lambda_1)$$
$$A + j(-B) = L_W(\lambda_2) + j\Big(-L_H(\lambda_2)\Big)$$
$$\text{their sum is } 2A = L_W(\lambda_1) + L_W(\lambda_2).$$

This shows that the hue component of one has been cancelled by the hue component of the other.

This theory has been experimentally checked. Sinden (2) took twelve pairs of complementary wave-lengths and experimented in a most careful way with them. He set up a match field of white,

whose luminosity was kept constant at a value of 75 photons. The task of the experimenter was to pick out pairs of monochromatic lights which were complementary and also to regulate their intensities so that he could get a perfect match of the field in question. Sinden, therefore, set for himself the double task of getting two hues that would cancel each other and that would sum up in intensity to 75 photons.

$$(1) \quad L_W(\lambda_1) + L_W(\lambda_2) = 75 \text{ photons}$$

$$(2) \quad L_H(\lambda_1) + L_H(\lambda_2) = 0$$

Under the conditions of this experiment we have two values which can be determined readily. One of them needs not be determined for it is given as one of the conditions, namely:

$$L_W(\lambda_1) + L_W(\lambda_2) = 75$$

The other condition to be determined experimentally is that

$$\frac{L_{\lambda_1}}{L_{\lambda_2}} = C$$

We have therefore these two simultaneous equations, which can be easily solved.

One matter then which comes up here that needs a little discussion is that of the units that are used. In all the equations so far given the units assumed are purely arbitrary depending upon the units that had been previously employed in setting up the primaries. The units used in the above equations may have been most anything. Under the conditions of Sinden's experiments the units are to be expressed in photons. This can be done in this way: Let us indicate with primes the *L, V, G, R,* etc., when they are transformed so as to express photons and keep them unprimed to express the relations in terms of our old units.

Since our resulting luminosity is to be 75 photons, we may assume that 35 of these units are the effect of λ_1 and that the remaining 40 are the result of the action of λ_2.

In terms of the old units we have

$$L_{(\lambda_1)} = V_{(\lambda_1)} + G_{(\lambda_1)} + R_{(\lambda_1)}$$

$$L_{(\lambda_2)} = V_{(\lambda_2)} + G_{(\lambda_2)} + R_{(\lambda_2)}$$

To transfer these luminosities into photons we have these relations:

$$kL_{(\lambda_1)} = 35 = L'_{(\lambda_1)}$$
$$kL_{(\lambda_2)} = 40 = L'_{(\lambda_2)}$$

and we express the hues as V′, G′, and R′ from which we get

$$L'_{(\lambda_1)} = V'_{\lambda_1} + G'_{\lambda_1} + R'_{\lambda_1}$$
$$L'_{(\lambda_2)} = V'_{\lambda_2} + G'_{\lambda_2} + R'_{\lambda_2}.$$

These last two equations express the hues as luminosity for us in terms of the photon as our unit.

A scrutiny of the following tables of Sinden's (2) experiment

TABLE 1

COMPOSITION OF COMPLEMENTARY COLORS IN TERMS OF PRIMARIES V, G, AND R

From Hecht (2)

Complementaries		Luminosities	
λ_1	λ_2	L_1	L_2
650	498	31.6	43.4
609	493.5	39.3	35.7
591	490	50.3	24.7
586	487.5	55.4	19.6
580	482.5	62.0	13.0
578.5	480.5	64.2	10.8
576.5	477.5	65.8	9.2
575.5	474.5	67.7	7.3
574	472	69.0	6.0
573	466.5	70.6	4.4
572	459	72.1	2.9
570.5	443	73.2	1.8

TABLE 2

From Hecht (2)

Complementaries		Luminosities		Composition		
λ_1	λ_2	L_1	L_2	V_1+V_2	G_1+G_2	R_1+R_2
650	498	31.6	43.4	23.2	25.8	26.0
609	493.5	39.3	35.7	24.3	26.3	24.3
591	490	50.3	24.7	24.3	26.2	24.7
586	487.5	55.4	19.6	24.3	26.0	24.7
580	482.5	62.0	13.0	24.5	25.8	24.7
578.5	480.5	64.2	10.8	24.5	25.8	24.7
576.5	477.5	65.8	9.2	24.7	25.6	24.7
575.5	474.5	67.7	7.3	24.7	25.6	24.7
574	472	69.0	6.0	24.7	25.6	24.7
573	466.5	70.6	4.4	24.7	25.4	24.7
572	459	72.1	2.9	24.9	25.4	24.7
570.5	443	73.2	1.8	24.7	25.6	24.7

shows that the pairs of wave-lengths he chose gave—as indicated under the column labelled *Composition*—practically equal numbers of receptors activated by each color, thus showing that the wave-lengths chosen were white complementary pairs.

Summary

Any wave-length initiates both the experience of hue and that of brightness. Luminosity is a property of hue and of whiteness (or grayness) and can be expressed as either a triplet or a doublet vector. The algebraic processes of addition and subtraction are applicable to both hue luminosity and the white luminosity of a given wave-length.

In the case of complementary pairs of wave-lengths it is shown that the real component of the vector expressing one of the pair is added to the real component of the other, whereas the imaginary components are of opposite signs and therefore cancel each other.

References

1. Allen, F. On the discovery of four transition points in the spectrum and the primary colour sensations. *Phil. Mag.,* 1926, **38**, 55-81.
2. Hecht, S. The development of Thomas Young's theory of color vision. *J. Opt. Soc. Amer.,* 1930, **20**, 231-270.
3. Sinden, R. H. Studies based on spectral complementaries. *J. Opt. Soc. Amer. & Rev. Scient. Instruments,* 1923, **7**, 1123-1154.

The Ohio State University
Columbus, Ohio

L'ALGÉBRE DE LA LUMINOSITÈ A L'ÉGARD DES PAIRES COMPLÉMENTAIRES BLANCHES DE LONGUEURS D'ONDE

(Résumé)

Toute longueur d'onde initie l'expérience de teinte et celle de clarté. La luminosité est une propriété de teinte et de couleur blanche (ou de couleur grise) et peut être exprimé ou comme vecteur triplé ou comme vecteur doublé. Les processus algébriques de l'addition et de la soustraction s'appliquent et à la luminosité de teinte et à la luminosité blanche d'une longueur d'onde donnée.

Dans le cas des paires complémentaires de longueurs d'onde on montre que la vraie composante du vecteur qui exprime l'une des paires est ajoutée à la vraie composante de l'autre, tandis que les composantes imaginaires sont de signes contraires et ainsi se cancellent.

Williams

DIE ALGEBRA DER LUMINOSITÄT, MIT BEZUG AUF WEISSE, EINANDER ERGÄNZENDE [COMPLEMENTARY] WELLENLÄNGENPAAREN [PAIRS OF WAVE-LENGTHS]

(Referat)

Jede Wellenlänge [wave-length] erzeugt sowohl eine Farben wie eine Helligkeitswahrnehmung [experience of brightness]. Luminosität ist eine Eigenschaft der Farbe [hue] einerseits und des Weiss oder Grau anderseits. Sie lässt sich entweder als einen dreifachen [triplet] oder als einen Doppel-Vektor deuten. Die algebraischen Vorgänge der Addierung und der Subtrahierung sind sowohl auf Farbenluminosität wie auf Weissheitsluminosität anwendbar. In Bezug auf einander-ergänzende [complementary] Wellenlängenpaare wird erwiesen, dass der aktuelle [real] Bestandteil des einen-der-beiden-Mitglieder-der-Paaresdarstellenden [expression] Vektors dem aktuellen Bestandteil des anderen hinzugefügt wird, während die imaginären Bestandteile entgegengesetzte Zeichen tragen [are of opposite signs] und folglich einander aufheben [cancel each other].

WILLIAMS

PSYCHOPHYSICS AND BEHAVIOR*

From the Psychological Laboratories of Clark University

C. H. Graham

The problem of descriptive correlation for a situation involving a responding human being and a variable stimulus intensity is a problem of long standing. Historically, it is bound up with the question of the measurement of sensation intensity (Titchener, 8). Functional relationships which were given as descriptions of correlations in typical "psychophysical" experiments (e.g., the Weber-Fechner law) provided the motivation for a voluminous controversial literature (cf. Titchener, 8, for an historical summary).

The opposition which developed against "psychophysics" was, in general, based on the logic of the distinction between mind and body. Mind, according to the opponents, is not possessed of magnitude and for this reason the measurement of sensation is impossible. This is the objection which has been called the "quantity objection" (Boring, 1), and the history of "psychophysics" is largely the history of an attempt logically to circumvent this difficulty. Boring (1) has enumerated five ways in which the difficulty has been met. He argues that, if we accept a description such as that given by the Weber-Fechner law, a relationship which involves two variables, an S and an R, which are covariant and not identical (the modes of variation being different), we may account for the difference in variation by the following devices: (1) We may argue, as Müller did and as Hoagland has done later (3), for a physiological interpretation of Weber's law. (2) We may introduce the concept of *Merklichkeitsgrade* (Wundt). (3) We may judge either stimuli or sensations (Fechner) and determine which is to be observed. (4) We may substitute sense distance for sensation magnitude (Ebbinghaus, Titchener, James, *et al.*). (5) We may restrict ourselves to the observation of stimuli and the errors incurred in the observation (Fullerton and Cattell).

The present discussion presents still another treatment of the intensity function. Specifically, it deals with an attempt to de-

*Accepted for publication by Carl Murchison of the Editorial Board.

scribe the function as a stimulus-response correlation. An analysis is made of the response variable in certain "psychophysical" experiments; the stimulus-response functions are described and characterized; and, finally, conclusions are drawn as to the nature of the inferences which may be made on the basis of these functions.

I

In any experiment with an intact organism the first step in description is the formulation of a relationship between a variable aspect of the stimulating conditions and some convenient measure of a response. [Cf. Skinner (7) on the problem of "measure of response."] The measurable feature of the response will depend for its magnitude upon a great many variable features of the stimulating situation. Thus we may say that

$$R = f\ (a, b, c, d \ldots\ldots) \qquad [1]$$

where R is the magnitude of the variable feature of the response and a, b, c, d, etc., are magnitudes of various stimulus variables. For any given experimental situation all but one of the variables, a, b, c, d, etc., are maintained constant and R is determined as a function of one independent variable. Expressed in the form of an equation,

$$R = f\ (a, b_1, c_1, d_1 \ldots\ldots) \qquad [2]$$

where R is the value of the dependent variable, a the value of the independent variable, and b_1, c_1, d_1 constant values of b, c, and d. Among those variables which are maintained constant is the variable of "instruction stimulus."

Stimulus-response correlations in a great number of behavior experiments may be described by this formulation. Thus, the latent period or magnitude of the pupillary reaction may be determined as a function of the light intensity with the variables of exposure time, time in the dark, area of stimulus surface, etc., maintained constant. Or we may determine the "error" aspect of a rat's response as a function of the number of presentations of the maze situation, with all other variables constant.

However, when we consider the complications inherent in the typical psychophysical experiment it might seem that here the direct stimulus-response formulation would be difficult of application. Nevertheless, it seems to the writer that the methods and results

of these experiments may be treated as special cases of the formulation in equation 1, and a demonstration is attempted in the following paragraphs. It would be difficult to consider all of the various psychophysical methods, and for this reason we shall restrict ourselves to a consideration of three which have been developed for the measurement of "sensation intensity," particularly as they apply to measurement in vision. It is hoped that a discussion of these three will carry implications for the treatment of other methods and variations of the methods under consideration.

The three methods which have been chosen are those given by Richardson (6), who says: "There are at least three ways of measuring sensations of hue, brightness or saturation as distinct from stimuli S: (E) by counting small equal appearing intervals; (J) by counting just perceptible intervals; (R) by directly estimating the *ratio* of unequal intervals both much larger than the least perceptible." Let us consider each of these methods separately.

Method R. It may be shown that the conditions and typical results of the R method may be fitted by a formulation such as that given by equation 2. In the concrete experimental situation the subject may respond to a specific intensity of stimulus by putting a mark on a scale marked zero and 100 at its two extremities. The aspect of the response measured is the distance from one of the two extremes. Another possible variation of the experiment would involve the measurement of the angle made by the forearm with the upper arm; or we might measure the height of a potential wave developed presumably by the activity of certain muscle groups. Perhaps one of these response aspects might have more correlation and test value than the others—but it is necessary to emphasize that the marking on a scale in terms of distance from an extremity of the scale is just as much an aspect of response as is latent period or tension development. The typical stimulus-response function given by this method [if we are allowed to conclude from results obtained by Maxwell (5) in experiments in which wave-length distribution rather than intensity was the independent variable] is continuous within limits, i.e., the response magnitude varies continuously with the intensity of the stimulus presumably between the limits of threshold intensity and the upper limit of intensity discrimination.

Method J. When we turn to methods E and J we find that

our attempt at a stimulus-response formulation is complicated by the fact that the responses observed in these methods are discontinuous functions of the stimulus. In order to elucidate this point let us consider for a moment the procedure involved in method J, the method used for the determination of the just-discriminative difference. For purposes of simplification we shall restrict the categories of response to two: "brighter" and "not brighter." (Thus we shall eliminate the categories of "equal" and "less bright," both being subsumed under the category "not brighter.") The subject is placed before a stimulus surface which is separated into two halves, each of the halves being independently illuminated. The illumination on one half (the comparison area) is maintained constant, while that on the other (the test area) is variable. The subject is instructed to report "brighter" when the test area is discriminated as more intensely illuminated than the comparison area. Under all other conditions his response is "not brighter." Thus the reaction to any condition of illumination must be one of a possible two. The task of the experimenter consists in noting at what illumination of the test area (with illumination on the comparison area constant) the response "brighter" is substituted for the response "not brighter" or vice versa.

If it be agreed that the above description is a fair (if schematic) characterization of the procedure involved in method J we may go on to a consideration of the problem: How do these measurements relate to the stimulus-response function of equation 2? In answering this question, let us first consider the problem of the change from a response of "brighter" to one of "not brighter." Below a critical value of intensity[1] and presumably down to the absolute threshold the response is a constant "not brighter." Above the critical value and presumably up to the limits of visual discrimination the response is a constant "brighter." In terms of such a description the response may be treated as a discontinuous function

[1]Here the objection might be raised that such a statement can be true only if by "critical value" we mean a *most likely critical value* which is determined by a statistical procedure. The objection is valid; and it is valid not only for the experiments of "psychophysics" but for the experiments of all physical science. Ordinarily, any experimentally determined "point" is in reality a "most likely point," and thus our correlations are functional expressions for relationships between "most likely values."

of one variable aspect of the stimulus situation, i.e., the intensity of illumination on the test area.

It is difficult to state an exact functional relation between stimulus and response in this situation because of the limitations imposed upon us by our system of units of visual measurement. Thus we have no method for assigning values to the auditorily received aspects of the responses "brighter" and "not brighter." However, the question of units is of little moment for this discussion. The statement that, within certain limits, the response is constant and that, within other limits, it is different but constant in this difference is sufficient to define the response as a discontinuous function of the stimulus variable.

If we wish to obtain a functional relation between the response and the stimulus we may assign a numerically constant value, p, to a response "brighter" and a numerically different constant value, q, to a response "not brighter." These two constants, p and q, are arbitrary except for the condition that they are unequal; their absolute values will depend upon the method of measurement and the aspect of the response measured. The inequality between the two constants expresses mathematically some difference between the measured response aspects. We may then draw a graph of the relation between response magnitude and intensity which will have the form of Figure 1. The function which is obtained is defined by the conditions that $R = q$ for $I < I_0$, and $R = p$ for $I > I_0$. Here I_0 represents the value of I which occurs at the discontinuity of the curve.

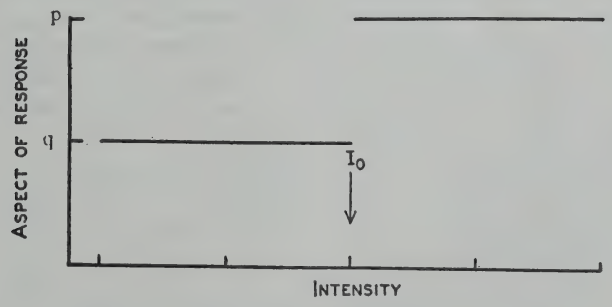


FIGURE 1

GRAPH OF A POSSIBLE STIMULUS-RESPONSE FUNCTION FOR METHOD J

Discontinuous functions of this type can, in certain cases, be expressed analytically in terms of known functions of mathematical analysis (e.g., by a Fourier integral). However, regardless of the expression used, it is important to realize that a function such as that shown in Figure 1 may be represented as a special case of equation 2. It should further be noted that the values assigned to the ordinate are not immediately important for this discussion. In elaborations of this experiment, where more variables are present than the variable of test-patch intensity, historically the dependent variable has been I_0, and this has been designated by any two or more responses which vary in a discontinuous manner with intensity. One very common elaboration results in the functional relation $R = f\ (a, b, c_1, d_1 \ldots .)$. This is the formulation required for the typical experiment on the verification of the Weber-Fechner law, where not only is intensity on the test area varied $(a \equiv I_t)$, but also intensity on the comparison area $(b \equiv I_c)$. ΔI may then be defined as $I_0 - I_c$, where I_0 is the value of I_t at the discontinuity of the curve. We might elaborate on situations requiring a greater number of variables, e.g., area, but this is a problem which cannot be treated adequately here.

Method E. One of the possible variations of method E as applied to a visual experiment is the following. The subject is presented with two illuminated patches, these patches differing in intensity of illumination. For any one experimental series the illumination on the patches is kept constant. Between the two constant illuminated areas is another stimulus surface on which intensity of illumination is variable. The subject is instructed to say "yes" when he discriminates a specific intensity of illumination on the middle area as half-way between the intensities on the outer areas. His response to all other intensities is to be "no." The experimenter's task is to note the intensity discriminated and to control the necessary variables in the stimulus situation.

In this experiment, as in the experiment by method J, we find that the response magnitude is a discontinuous function of intensity. Within certain limits of intensity, the response is "no." At a specific intensity the response changes to "yes" and remains "yes" for a small interval of intensities (the just-discriminable difference), thereafter becoming "no" again. If we assign arbitrary expressions to these responses as we did for the responses of method J, our

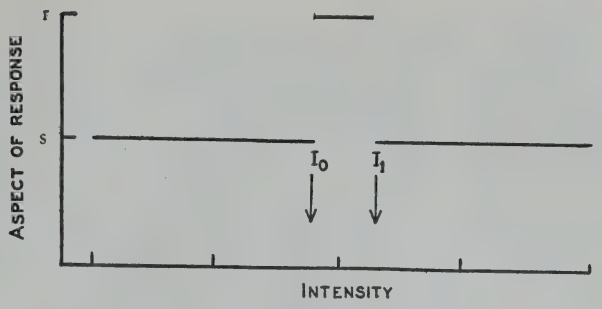


FIGURE 2

GRAPH OF A POSSIBLE STIMULUS-RESPONSE FUNCTION FOR METHOD E

function is as represented in Figure 2, where some aspect of the "yes" response has been assigned a value of r and some aspects of the "no" response a value of s. The function thus plotted may be described by the condition that $R = s$ for $I < I_0$ and $I > I_1$, and $R = r$ for $I > I_0$ but $< I_1$. Such a function may be expressed as a known function of mathematical analysis.[2]

II

The treatment of the three psychophysical methods as given above makes it possible to define stimulus-response functions in certain variants of the "sensory" experiment. Applications of these formulations to other aspects of "sensory" experimentation may be made without difficulty, e.g., in determinations of threshold (where variables due to a comparison area are absent).

Let us now continue with the examination of the discontinuous stimulus-response functions. In this examination we are almost immediately confronted with the problem: What is the nature of the response aspect measured at the discontinuities of the graphs in Figures 1 and 2?

In the discussion of methods E and J it was found that at the critical value of intensity the measured value of the response changes to a new constant value. Only a bit of reflection is required to show that, no matter what other variables of the stimulating situa-

[2]If we disregard the just-discriminable difference the interval I_0I_1 approaches a point, I_2, as a limit and the function is defined by the condition that $R = r$ for $I = I_2$ and $R = s$ for $I \neq I_2$.

tion we change (e.g., intensity of comparison area, area of either patch, time of exposure, etc.), the determination of the critical value of intensity is always made on the basis of the change to a new response value. Further, this change is a constant change for it involves only the substitution of the response "brighter" for "not brighter" (or "yes" for "no"). In determining, then, the critical value of intensity we really determine the intensity of stimulus which is required to make the subject respond with a response at least one aspect of which is constant throughout the series of determinations. *Thus, when we determine how this critical value varies as a function of other variables of the stimulating situation,* e.g., the intensity of the comparison area, *we must regard our ultimate description as indicating the manner in which the stimulus variables vary for the production of a constant behavior effect.*

As a specific example for the elucidation of this point we might consider the experiments of König and Brodhun (4), which were made to test Weber's law for vision. In these experiments the response categories were presumably "less" and "equal." The final description of the results of this experiment is embodied in a graph in which $\Delta I/I$ is plotted as a function of log I, ΔI representing the difference in intensity between the intensity of the comparison area and the intensity of the test patch at the response discontinuity, and I being the intensity of the test area at the response discontinuity. This graph, then, shows the manner in which certain functions of the stimulating variables vary; and it is not until we examine the conditions of variation that we gain an insight into the response side of the description. The full description of these findings requires the statement that the relationship given by the curve is the relationship obtained with the variables in such combination as to produce a constant response. Looked at in this light the function as given by the graph is a stimulus-response function; it describes the interrelationships of variables with reference to a constant effect. Ultimately the description of functions of this type may require a definite expression of the magnitude of the constant effect. This would be necessary if it were shown that the experimental values varied as a function of the magnitude of the constant effect (in some significant units). Such a possibility seems quite probable.

III

So far in this discussion we have been concerned only with a formulation in terms of behavior. However, it may be that we shall not be satisfied to restrict our discussion to this level, i.e., we might like to consider the events intervening between the stimulus and response.

Let us consider the steps involved in an analysis of this type. The next step beyond the descriptive one of expressing the response as a function of the stimulus is one of hypothesis. Thus when attempting to account physiologically for the events intervening between a stimulus and a response it is our practice to set up hypotheses in terms of mechanisms which we already know. By following this procedure we may set up an hypothesis as to the rôle of impulse frequency in the afferent fibers of an intact organism; or we might, in the case of the eye, formulate a hypothesis in terms of what we know of photochemical processes. However, regardless of the hypothesis which we may adopt, the value of the hypothesis will depend upon the extent to which it may be tested and upon the manner in which it meets these tests. More specifically we may say that the value of the hypothesis depends upon the extent to which, when applied to a new situation in which the variables of the original formulation assume new values, the original formulation may still serve (by the logical consideration of the differences in value) to give an account of the proper interrelationships of the variables in the new situation. Conversely, the interrelationships of the variables as observed in the new situation must allow us to recover, by a logical procedure, the interrelationships of the variables in the original situation. Often, as a result of the test, it becomes necessary to modify the hypothesis in order to describe the contributions of new variables—variables which had been overlooked or which had entered as constants into the original hypothesis.

Ultimately, the hypothesis may become a theory. This is its status when it has been shown by test to be compatible with our observations and logical conclusions concerning the correlations which it presumes to describe. As an account of phenomena, the theory, in this state, presents us with what we must accept as a true representation. For all this the theory is an inference; and when we attempt to describe the events intervening between the stimulus and

the response, these events, for the intact organism, must be a matter of inference.

And now we may ask: What may be the nature of our inferences as to the processes intervening between stimulus and response? In the present state of our knowledge concerning the structure and functions of organisms our theories of these intervening processes may be of three kinds: (1) We may set up an account in terms of physiological processes; (2) we may discuss the characteristics of the conscious states of the subject; or (3) we may hypothesize in terms of both physiological processes and mental states.

In the description of any functional relationship the problem of units of measurement is of fundamental significance. The most immediately obvious advantage of the account in terms of physiological processes is that it describes in terms of concepts which are common to the general body of scientific knowledge; that is, it systematizes our experienced reality in terms of magnitudes derived from the concepts of space, mass, and time. Thus, our generalizations as to the structure and function of cells, in so far as these generalizations involve functional correlations which are the results of operations of measurement, are directly or indirectly referable in their units of expression to the fundamental units of space, mass, and time.

The hypothesis in terms of conscious states involves the inclusion of events which are not amenable to characterization in terms of these fundamental systematic concepts of science, i.e., events which are not describable in terms of units of space, mass, or time. Specifically, this means that, unless we define consciousness, as Boring (2) has done, in terms of these concepts, we must describe the states of consciousness in terms of units other than those of space, mass, and time. But here, from the point of view of scientific method, we meet with a difficulty in that consciousness is inferred from our measurements which are given in the accepted units. By the methods of science we must ultimately state the hypothesis in terms of such functional relationships that, in the proper logical combination, these relationships will allow us to recover the observed relationship. The relationships, given in symbols which are magnitude expressions, are of significance as indices to time, mass, space, and their derivative aspects of reality. If a different kind of event, not involving any of these concepts, enters into a sequence started

by an initiator which is describable in these concepts and ends with an effect describable in these concepts we shall have to assume (*a*) that any event not describable in these concepts is somehow cancelled out in our hypothesized intermediate steps or (*b*) does not enter into these steps. Since either description is hypothetical we shall have to choose the one which may be tested. Obviously, we have no means at hand for testing the first possibility since by definition the event is different in kind from the events which may be used for test purposes. In these circumstances we should never be able to recover our observed relation from a consideration of the intervening conscious process. Because of this failure it seems to the writer that a consideration in terms of conscious states is irrelevant to scientific description. On this basis it would seem that an adequate account of data in the so-called "psychophysical" experiment may be given on the level of observation in terms of stimulus and response, and on the level of hypothesis in terms of physiological correlation.

Summary

1. Stimulus-response functions are described for three kinds of "psychophysical" experiment.

2. There is a discussion of the response variable in certain of these experiments.

3. The problem of description and explanation is considered as it applies to the "psychophysical" experiment.

References

1. Boring, E. G. The stimulus-error. *Amer. J. Psychol.*, 1921, **32**, 449-471.
2. ————. The physical dimensions of consciousness. New York: Century, 1933. Pp. xii+251.
3. Hoagland, H. The Weber-Fechner law and the all-or-none theory. *J. Gen. Psychol.*, 1930, **3**, 351-373.
4. König, A., & Brodhun, E. Experimentelle Untersuchungen über die psychophysische Fundamentalformel in Bezug auf den Gesichtsinn. *Sitzber. d. Akad. d. Wiss.*, Berlin, 1889, 447-450.
5. Maxwell, R. S. The quantitative estimation of the sensation of colour. *Brit. J. Psychol.*, 1929-1930, **20**, 181-189.
6. Richardson, L. F. The measurability of sensations of hue, brightness or saturation. In *Report of a joint discussion on vision held on June 3, 1932, at the Imperial College of Science by the Physical and Optical Societies.* London: Physical Soc., 1932. Pp. 112-116.

7. SKINNER, B. F. The concept of the reflex in the description of behavior. *J. Gen. Psychol.*, 1931, **5**, 427-458.
8. TITCHENER, E. B. Experimental psychology: a manual of laboratory practice. Vol. II. Quantitative experiments. Pt. 2. Instructor's manual. New York: Macmillan, 1905. Pp. clxxi+453.

Clark University
Worcester, Massachusetts

LA PSYCHO-PHYSIQUE ET LE COMPORTEMENT

(Résumé)

1. On décrit les fonctions stimulus-réponse pour trois sortes d'expériences "psycho-physiques."
2. On discute la variable des réponses dans certaines de ces expériences.
3. On considère le problème de description et d'explication comme il s'applique à l'expérience "psycho-physique."

GRAHAM

PSYCHOPHYSIK UND VERHALTEN [BEHAVIOR]

(Referat)

1. Es werden Reiz-Reaktionsfunktionen [stimulus-response functions] für drei Arten von "psychophysischen" Versuchen beschrieben.
2. Es wird die Reaktionsvariable bei einigen dieser Versuche besprochen.
3. Es wird die Aufgabe der Beschreibung und Erklärung in Bezug auf das psychophysische Experiment berücksichtigt.

GRAHAM

SPATIAL SUMMATION OF SUBLIMINAL STIMULI IN THE RETINA OF THE HUMAN EYE*[1]

From the Psychological Laboratories of Clark University

ROBERT J. BEITEL, JR.

INTRODUCTION

The problem of "spatial effects" has long been a recognized problem in human vision. The observation (*a*) that a variation in the size of the retinal area stimulated has an effect similar to that produced by a corresponding variation in light intensity, and the observation (*b*) that the effect obtained by stimulation of one area of the retina is, to some extent, dependent upon the nature of the stimulation of adjacent areas have served as points of departure for a great many experimental attacks. Many of the early investigators (Riccò, 18; Abney, 1; Piper, 16; Abney and Watson, 2; and others) concerned themselves primarily with empirical quantitative formulations of the various effects demonstrable at the threshold. Later work by Kleitman and Piéron (14) has served to discourage further attempts to obtain a generalized formulation for the area-intensity function, for, as these authors point out, it is unlikely, because of the "heterogeneity of the retina," that the "area effect" can be described in terms of a simple law. It was not until the work of Adrian and Matthews, (3, 4) that a concrete explanation of spatial effects was made possible. These investigators found, in recording impulses from the optic nerve of the Conger eel, that the latent period of discharge decreases when the retinal image is increased in area and when the number of separately illuminated areas is increased. These results are interpreted as demonstrating the presence of summation in the retinal synapses, an interpretation which is strengthened by observations on the effect of strychnine. Granit (11) has shown more recently that the typical area effects are reflected in the behavior of the

*Recommended by C. H. Graham and accepted for publication by Carl Murchison of the Editorial Board.

[1]The author wishes to express his indebtedness to Dr. C. H. Graham, under whose direction this research was conducted.

retinal potential of the cat, i.e., by a decreased latent period with increased area; and Graham (8) has furthered the analysis of the locus of spatial effects by the demonstration that, in an eye which lacks retinal synaptic connections, the latent period for nerve discharge is unchanged by an increase in the number of fibers activated.

During the last few years Granit and his co-workers (Granit, 10; Granit and Harper, 13; Granit and Hammond, 12; Graham and Granit, 9) have succeeded in demonstrating the importance of spatial effects in the human retina at high intensities of illumination. These researches have involved the use of flicker as an indication of retinal excitation; and results, particularly by Granit and later by Granit and Harper, may be interpreted as showing functions of the human retina which parallel those of the excised preparation. More evidence on this parallelism is present in two reports by Geldard (6, 7).

Certainly it is true that the human retina is morphologically similar to the retinae of other vertebrates used for action-potential determinations. For this reason it would seem that effects determined in the excised eye-nerve preparation should be demonstrable within limits by the technique of sensory experimentation. Further, it would seem that, if these effects were demonstrable, it should be possible to carry on further researches on the synaptic basis of visual reception and ultimately to formulate a description of these processes. Such a program might well be productive of valuable facts and inferences for our knowledge of central visual mechanisms.

The present paper, in line with such an approach, is concerned with a description of a typical synaptic process which seems to play a rôle in the retina similar to that which it plays in the spinal cord. The activity demonstrated is the summation of subliminal excitatory effects. Further analysis of the phenomenon seems to present significant material for our understanding of the nature of the threshold.

The method which has been used is the following: The threshold for two areas separated by varying spatial intervals is determined, as is the threshold for either area alone. The interpretation in terms of the spatial summation of subliminal stimuli seems warranted by the finding that, within limits of spatial separation, the threshold for the two areas is lower than that for either one alone.

Apparatus and Procedure: Peripheral Determinations

The apparatus and procedure are of the conventional type used for determinations of the threshold.

A 250-watt Mazda daylight lamp (maintained at 110 volts d. c.) enclosed in a light-tight box, together with a cooling-cell and a biconvex lens, provides the source of illumination. The rays of light coming from the lamp pass through the cell and lens and are brought sharply to a focus in the plane of the slit in a pendulum. The pendulum has a large, heavy, iron frame; it is of sturdy construction and is balanced upon steel knife-edges. An aluminum plate, containing a slit, is attached to the top of the upper shaft. The exposure time is determined by the size of this slit and by the speed at which it passes through the focus of the illuminating rays. Light intensity is controlled by means of a neutral-tint wedge (adjustable by rack and pinion) and a balancing-wedge placed in front of the aperture of the light-box. The slit-aperture in the plate of the pendulum permits exposure of the light beam in flashes having a duration of 15 sigma when the pendulum is in motion. Directly behind the slit-aperture and attached to the aluminum plate is a holder for Wratten neutral-tint filters. The rays of light which pass through this system diverge and illuminate a milk-glass plate which is set in a large wall-board screen. Cards bearing the stimulus patches can be slipped into place in a container directly behind the milk-glass plate. A small box containing a flashlight bulb which faintly illuminates a red cross cut in the surface of the box provides a fixation target. This target is attached to the screen on a level with, and 15 degrees peripheral to, the central fixed stimulus patch. A chair and headrest are provided for the subject, who is seated at a distance of 170 cm. from the screen.

Seven cards bearing stimulus patches were used in the experiment. Each card contained two semicircular apertures of 0.5-degree radius. In successive arrangements the distance between the semicircles varied from no separation, i.e., a complete circle, to a separation (in terms of visual angle) of 150 minutes. The semicircles were so arranged that one of the stimulus patches remained in the same constant position, the other being moved toward the periphery. These patches are designated as the "fixed" and "variable" patches, respectively. The patches could be illuminated simultaneously or singly.

Dark-adaptation was maintained for from 40 to 45 minutes before the first threshold determination. After this interval the subject was instructed to direct regard with the right eye to the red cross; vision with the left eye was occluded in all the experiments. In order to insure adequate fixation, the subject was instructed to signal the experimenter when the cross "appeared clear in outline." Upon this signal the experimenter released the pendulum, permitting exposure of one flash. Sufficient time (at least 30 seconds) was allowed to elapse between exposures in order to obviate anomalous visual effects. During these periods, vision to both eyes was occluded or the fixation light was turned off. Results obtained under these conditioned were found to be consistent over a period of two hours.

All of the experimentation was carried on at approximately the same time of day. Determinations were usually made on three cards at each sitting. The order of presentation of the cards varied from subject to subject and in different series with the same subject.

The procedure was the same for each measurement: First the threshold was determined for both semicircles together, then for each alone; and finally for both together again.

The method for determining the threshold intensity requires some description. Obviously, in measurements of this kind we are not interested in determining a representative value for the subject's discriminative behavior. Rather, we are interested in getting a series of determinations in a short time—during which time the subject is presumably in a stable physiological state, and during which time his threshold values are highly reproducible.

In order to fulfill this requirement the method of limits was used. The subject's response was tested five times with an ascending series and five times with a descending series over a range of wedge readings which were shown by preliminary experiments to embrace the range of uncertainty. The intensity below which the response was always "no" was recorded, as was the intensity above which the response was always "yes" [i.e., "the light was seen"]. The range of the resulting interval of uncertainty was then plotted for each determination.

Results of the Peripheral Determinations

The averaged results for four subjects are given in Table 1 and Figure 1. These averages are obtained from the results of eight series, two for each subject.

TABLE 1

AVERAGED THRESHOLD INTENSITIES FOR FOUR SUBJECTS WITH THE "FIXED" PATCH AT 15 DEGREES IN THE PERIPHERY

Separation	Intensity (millilamberts) Both patches	"Fixed"	"Variable"
150′	.000089 ml.	.000089 ml.	.000101 ml.
95′	.000069	.000089	.000092
60′	.000064	.000092	.000095
19′	.000060	.000091	.000092
9′	.000055	.000088	.000095
5′	.000054	.000095	.000101
0	.000050	.000089	.000093

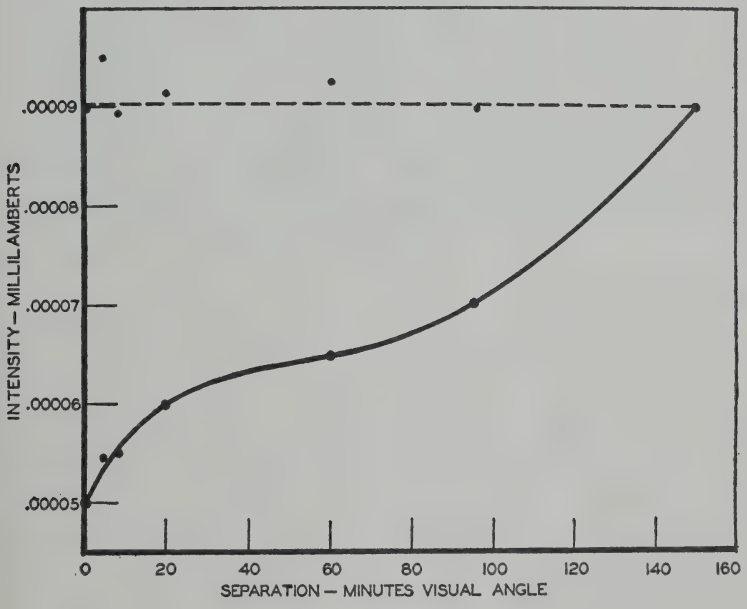


FIGURE 1

CURVE OF AVERAGED THRESHOLD INTENSITIES OBTAINED FOR BOTH PATCHES WHEN EXPOSED SIMULTANEOUSLY AT THE SEPARATIONS INDICATED

The dotted line is drawn through the points corresponding to the averaged values of the single patch having the lower threshold value at each particular setting. The "fixed" patch is at 15 degrees in the periphery.

Table 1 shows the averaged threshold intensities obtained for four subjects under the various experimental conditions. The column headed "Both patches" gives the threshold intensity for both patches together at the angles of separation shown in the preceding column. The column marked "Fixed" gives the threshold intensity for the semicircle which was at a constant angle of 15 degrees from the fovea; and the last column gives the threshold intensity of the patch which was placed at varying distances from the fixed patch.

Figure 1 gives a representation of the manner in which the threshold for both patches decreases as the distance between the patches decreases. At a separation of 150 minutes the threshold for both patches is the same as the threshold for either of the single patches—within the limits of a slight degree of variation. Thus when we plot the threshold values of the single patches on this graph we find that curve to be very nearly a straight line across the figure at an intensity value corresponding to the value of the widest separation. The dotted line in Figure 1 gives such a plot; the values of the points correspond to those of the single patch having the lower threshold value at each particular setting. The difference, then, between these values and the values as plotted for both patches together provides some measure of the summation due to the presence of the second semicircle.

A survey of this graph makes it clear that the range of variation in intensity is small—a range of 1 to 2, approximately. On a conventional logarithmic scale of visual intensity functions such a range is small in comparison with the total possible range. Consequently, it is essential that the greatest precautions be taken in order to insure the reliability of variations within such small limits. For this reason, then, it becomes necessary to consider the variability of the data. The reliability of the differences between the thresholds for the different separations may be ascertained by an examination of Figure 2 where the interval of uncertainty for each determination is plotted for a subject whose total range of intensity variation was smaller than that of any of the other four subjects. The results of this subject, obtained six months after the determinations on the other four subjects, are not included in the averages of Table 1. The inclusion would affect the values little. Although the curve for this subject is the most atypical, it nevertheless possesses the characteristics displayed by the curves for the other subjects: the initial sharp rise,

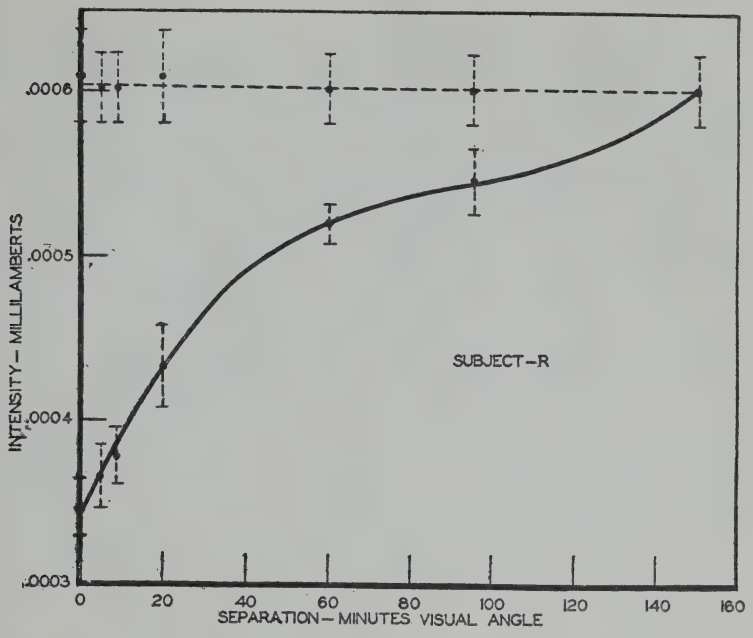


FIGURE 2

CURVE FOR ONE SUBJECT SHOWING THE RANGE OF THE INTERVAL OF UNCERTAINTY (SEE TEXT) FOR EACH DETERMINATION

The "fixed" patch is at 15 degrees in the periphery.

the intermediate interval of comparatively little change, and the final upswing to the point for the widest separation. The interval of uncertainty as described in the preceding section is the intensity interval between the value above which the response is always "yes" and the value below which the response is always "no." With the exception of the determinations made for the separation distance of 150 minutes, the absence of any overlapping between the ranges for the single patches and those for both patches together is quite evident. Furthermore, in the curve for both patches together the overlapping which occurs is found only in the cases of determinations for any two successive separations. Clearly the differences between the intervals of uncertainty for the greatest and smallest separations are seen to be reliable. On this basis it seems that there can be little doubt as to the reality of the summation effect.

Apparatus and Procedure: Foveal Determinations

The difficulties encountered in securing adequate fixation for foveal stimulation necessitated modifications of both apparatus and procedure. These difficulties were peculiar to the type of method tried and took such forms as extreme fatigue, accommodation and convergence variations, and "autokinetic phenomena." The following methods were employed in the preliminary investigations and were found to be unsatisfactory for our purposes (although it is probable that in other types of work, where the range of effect is large, they would be perfectly satisfactory): the method of blind-spot fixation; the method of central fixation utilizing peripheral guiding clues in the form of luminous paint spots; and a third method whereby a fixation-spot was projected upon one eye on a "point" corresponding to the area stimulated in the other eye. Since the results obtained under these conditions rendered interpretation hazardous, it was decided to employ a method whereby the determinations of difference thresholds, rather than absolute thresholds, could be used.

The apparatus was so arranged as to permit a dim, diffuse, illumination of the cards bearing the stimulus patches. The source of this illumination consisted of two flashlight bulbs enclosed in a small box, one side of which consisted of a milk-glass plate, set at an angle in front of the stimulus card. In the center of each card, and slightly below (5 minutes, in terms of visual angle) the level of the stimulus patches, was placed a small wedge of light gray paper which served as fixation object. The intensity of illumination on the gray fixation-spot required to make it just discernible at a distance of 4 meters (the fixation distance) was determined by preliminary test following a 5-minute period of dark-adaptation and was maintained constant throughout the experimental period. The higher light intensity necessary for stimulation under these conditions was obtained by interposing a large condensing lens in the path of the diverging rays between the pendulum and the milk-glass plate.

Five cards bearing stimulus patches were used in the experiment. Each card contained two square apertures, the sides of each square subtending a visual angle of 5 minutes. In successive arrangements the distance between the squares varied from no separation, i.e., a complete rectangle the longer side of which subtended a visual angle of 10 minutes, to a separation (in terms of visual angle) of 30 min-

utes. The squares were so arranged that the fixation-object always occupied a position midway and slightly below the level of the patches. The patches could be illuminated simultaneously or singly.

Each subject was first light-adapted for 2 minutes, then dark-adapted for 5 minutes, prior to the determination of each threshold. During these periods, and throughout the entire experimental period, the illumination of the stimulus card (and its fixation-object) was maintained constant. Light-adaptation was secured by having the subject regard a brightly illuminated surface (a milk-glass screen illuminated from behind) from a distance of 2 feet. The intensity of this surface was constant at 420 millilamberts.

The method of determining the threshold values was essentially the same as that used for the peripheral determinations and determinations were made in both directions for the entire series. It was found possible, however, due to the condition of constant, slight light-adaptation, to present the stimuli in more rapid succession and thus to obtain readings for a complete series in one experimental period. Binocular fixation was used for all the determinations.

Results of the Foveal Determinations

The averaged results for four subjects are given in Table 2 and Figure 3. These averages are obtained from the results of eight series, two for each subject. Table 2 presents the averaged threshold intensities obtained for the four subjects under the various experimental conditions. Figure 3 shows that, within the range of no separation to a separation of 10 minutes, the threshold intensity decreases as the distance between the two patches decreases. Within

TABLE 2

Averaged Threshold Intensities for Four Subjects: Foveal Values

Separation	Intensity (millilamberts)		
	Both patches	Single (A)*	Single (B)*
30′	.152 ml.	.166 ml.	.152 ml.
10′	.152	.155	.155
5′	.114	.157	.159
3′	.099	.159	.159
0	.088	.169	.159

**Single* (*A*) refers to the patch which appeared always to the right of the subject, and *Single* (*B*) to the patch which appeared always to the subject's left.

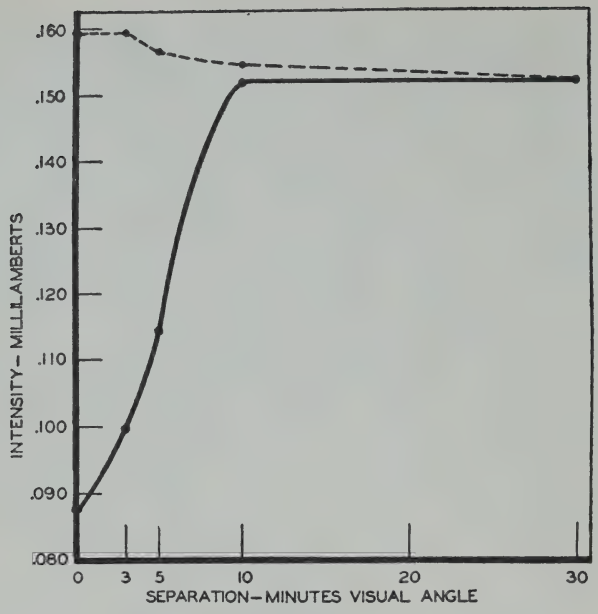


FIGURE 3

CURVE OF AVERAGED THRESHOLD INTENSITIES OBTAINED FOR BOTH PATCHES WHEN EXPOSED SIMULTANEOUSLY AT THE SEPARATIONS INDICATED

The dotted line is drawn through the points corresponding to the averaged values of the single patch having the lower threshold value at each particular setting. The determinations are made for the fovea.

the range of 10 minutes to 30 minutes of separation no change in threshold intensity is found. The dotted curve is drawn through the points corresponding to the values of the single patch having the lower threshold value at each particular setting. The difference between these latter values and the values as plotted for the two patches together provides some measure of the summation due to the presence of the second patch.[2]

In Figure 4 the range of the resulting interval of uncertainty is shown for each separation used with one subject whose responses

[2]The fixation object may, of course, give rise to an added excitatory component. However, this component is always constant and enters into the threshold for each patch when presented alone. Thus the summation noted as a function of the distance between the two variable patches must be a true differential effect.

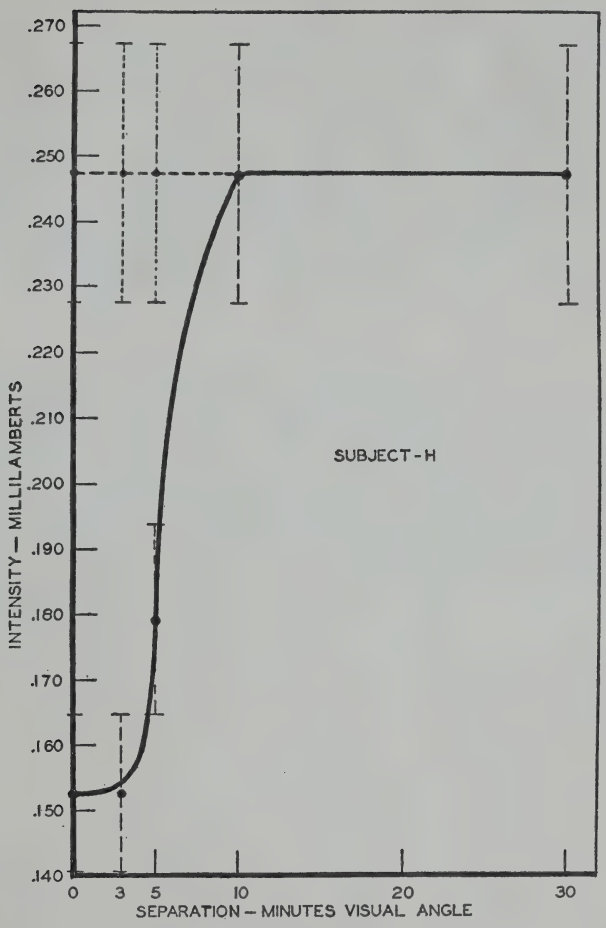


FIGURE 4

CURVE FOR ONE SUBJECT SHOWING THE RANGE OF THE INTERVAL OF UNCERTAINTY (SEE TEXT) FOR EACH DETERMINATION

The determinations are made for the fovea.

were quite consistent. (The results for the other three subjects are similar, differing only in that their values for the single patches are slightly more variable.) The dotted line, again, is representative of the threshold values corresponding to the single patch having the

lower threshold value at each particular setting. Thus it is seen that the differences between the thresholds for the different separations are quite significant and that there can be little doubt as to the reality of the summation effect.

It may be added at this point that, in the preliminary experiments, it was noticed in several cases that the threshold for both patches tends to drop again if separations greater than 30 minutes (45 minutes and 60 minutes) are used. This may be tentatively attributed to the function of incoming rods and is a point worthy of further investigation.

Discussion

The results of these experiments show that there is evidence for interaction between spatially separated, subliminally stimulated areas in both the periphery and fovea of the retina. The amount of this interaction is a function of the distance separating the stimulated areas, the effect increasing with increased proximity of these areas. That the process occurs over a limited range in both the fovea and the periphery is shown experimentally by the facts that, in the periphery, when the patches are separated by a visual angle of 2½ degrees, and in the fovea, when they are separated by a visual angle of 10 minutes, no effect can be noted. Under these circumstances the threshold intensity for both patches is the same as that of the single patch having the lower value. These results, in so far as they demonstrate interaction of a summative nature, are in line with those obtained by Adrian and Matthews (3, 4), Granit (10, 11), and Granit and Harper (13) in their studies on spatial summation where supraliminal stimuli were employed (see "Introduction").

The decrease in threshold intensity with a decrease in separation of stimuli is most easily interpreted as being due to an increased lateral summation which arises with an increased proximity of converging excitatory paths. Arguments for such an interpretation have been advanced at great length by the authors previously cited, and there is little reason for further discussion of this point.

In the peripheral retina convergence is exhibited in three forms: (1) convergence of many receptors upon one bipolar cell; (2) convergence of two or more bipolars upon one ganglion cell; and (3) the convergence of lateral internuncial neurons at a common focus. It is clear that the convergence of receptors upon a single bipolar cell

cannot explain a summation which occurs over a spatial interval of 150 minutes of visual angle. Nor can the convergence of bipolars upon ganglia explain the complete effect—although it is probable that the initial sharp rise in our curves may be correlated with these influences. The threshold, then, must be determined to some extent by the amount of interaction taking place over the lateral internuncial neurons. In this connection it is interesting to correlate our result, which shows summation over a visual angle of 2½ degrees, with the anatomical finding of Ramon y Cajal (17) that the limiting length of the lateral interconnections of the eyes of higher vertebrates is .8 mm. (corresponding to approximately 3 degrees of visual angle).

Similarly, if we consider the anatomical evidence, it may be possible for us to account for the initial sharp rise in the curves of Figures 1 and 2. Figures given by Chiewitz (5) on the relation of outer nuclei to ganglion cells show that at approximately 15 to 16 degrees in the periphery the ratio of outer nuclei to ganglia is about 11 to 1, and at approximately 20 to 22 degrees in the periphery this ratio increases to about 42 to 1. [At the time this work was done (1889) the chrom-silver method of Golgi-Cajal that enabled differentiation between the cells with regard to their connections was not in use; thus the outer nuclei need not all be the nuclei of rods and cones. Chiewitz's value for the fovea, however, which agrees with Koelliker's (15), shows a ratio of 1.09 to 1, and seems to indicate a fair degree of accuracy in the work.] In Figure 1 it is seen that the initial sharp rise occurs between zero separation and 20 minutes of separation. The retinal distance subtending a visual angle of 20 minutes is roughly about .09 mm. Peripheral rods and cones, it is estimated, have diameters of about .006 mm. and .003 mm., respectively. Consequently, the number of rods or cones that may exist side by side within a distance of .09 mm. may vary, roughly, from about 15 to 40. These figures seem to be in agreement with those of Chiewitz for the convergence of outer nuclei upon a single ganglion cell. Because of this agreement it might be reasonable to assume that spatial summation up to a separation of 20 minutes is largely mediated by the convergence of outer nuclei upon ganglia. The sharp rise occurring in Figure 1 between zero separation and 20 minutes of separation we attribute to this factor. Summation occurring over greater spatial intervals (up to 2½ degrees) will have

to be explained, then, in terms of convergence afforded by lateral internuncial neurons.

In the fovea, as is well known, anatomical convergence is quite strictly limited. For the most part evidence points toward a 1-to-1 relation between receptors and ganglion cells, but bipolars, amacrines, and some horizontal cells have been demonstrated (Ramon y Cajal, 17; Koelliker, 15). The number of ganglion cells increases greatly from the center of the foveal pit to the margin of the fovea, as do the bipolars.

Our results for the fovea, which exhibit summation within the limited range of 10 minutes of visual angle, correlate with this evidence. In contrast with the periphery, where converging pathways provide a basis for summation over approximately 2½ degrees of visual angle, the fovea exhibits little in the way of converging paths. Thus, we should expect that summation would be slight for any but the smallest of separations. On the other hand, the evidence indicates that within the range of the limiting small separation (10 minutes) summation does occur. As a matter of fact, the absolute range of the intensity effect is approximately the same in the fovea as in the periphery, i.e., the ratio of the threshold intensity with no separation to the intensity at the limiting separation is, in both cases, very nearly 1 to 2. The mechanism for this summation in the fovea may well be accounted for in terms of a spatially restricted, "direct" type of convergence—a type which would seem to be in harmony with our knowledge of foveal anatomy.

The appearance of threshold interaction has an important implication for interpretations which deal with the nature of the threshold. From the interpretations given above it follows that, in order to summate with another subliminal stimulus, the excitation due to one subliminal stimulus must reach a region of the nervous structure upon which the pathways from the two stimulated areas converge. Further, the results show that the excitation due to a subliminal stimulus may reach this region without causing a light-discriminatory response on the part of the subject. Such a response does not occur unless the excitation due to the stimulation of one region is reinforced by the excitation due to the stimulation of another region. This type of process is, of course, a nervous process and, by its nature, may occur in any part of the retino-cerebral system. However, the evidence at present is of such a nature as to indicate a retinal locus for

the effect. Thus, the experiments of Adrian and Matthews (3, 4) and Granit (11) show that functions of a similar nature are demonstrable in the excised eye and in decerebrate preparations. Obviously, in the case of the the excised eye, the only nervous structures which could account for the effects are the retinal layers. Since these observations on "preparations" are directly comparable with the results reported in the present experiment, it would seem that, in the present instance, it would be well to consider the effects as due to the activity of the retinal synapses. In these terms, the final response is judged to be initiated and controlled by the processes in the retinal layers and ganglia.

Summary

Threshold intensities have been determined for various separations of two stimulus patches in the fovea and periphery of the human eye.

The results may be summarized as follows:

1. In the periphery the threshold for both patches when exposed simultaneously decreases with decreased separation of these patches (within the limits of 2½ degrees of visual angle).

2. In the periphery the threshold for both patches when separated by a distance of 2½ degrees of visual angle is the same as the threshold of the single patch having the lower value.

3. In the fovea the threshold for both patches when exposed simultaneously decreases with decreased separation of these patches (within the limits of 10 minutes of visual angle).

4. In the fovea the threshold for both patches when separated by a distance of 10 minutes of visual angle is the same as the threshold of the single patch having the lower value.

These results are interpreted as demonstrating interaction of a summative type between subliminally stimulated areas. The amount of this interaction is a function of the distance separating the areas. Arguments are presented for the conclusion that the locus of the visual threshold effect is in the synaptic layers of the retina.

References

1. Abney, W. The sensitiveness of the retina to light and colour. *Phil. Trans. Roy. Soc.,* 1897, **190A**, 155-195.
2. Abney, W., & Watson, W. The threshold of vision for different coloured lights. *Phil. Trans. Roy. Soc.,* 1916, **216A**, 91-128.
3. Adrian, E. D., & Matthews, R. The action of light on the eye: I. *J. Physiol.,* 1928, **63**, 378-414.
4. ————. The action of light on the eye: II. *J. Physiol.,* 1928, **65**, 273-298.
5. Chiewitz, J. H. *Arch. f. Anat. u. Entwickelungsgesch.,* 1889, Suppl. Bd., 139.
6. Geldard, F. A. Brightness contrast and Heymans' law. *J. Gen. Psychol.,* 1931, **5**, 191-206.
7. ————. Foveal sensitivity as influenced by peripheral stimulation. *J. Gen. Psychol.,* 1932, **7**, 185-189.
8. Graham, C. H. The relation of nerve response and retinal potential to number of sense cells illuminated in an eye lacking lateral connections. *J. Cell. & Comp. Physiol.,* 1932, **2**, 295-310.
9. Graham, C. H., & Granit, R. Comparative studies on the peripheral and central retina: VI. *Amer. J. Physiol.,* 1931, **98**, 664-673.
10. Granit, R. Comparative studies on the peripheral and central retina: I. *Amer. J. Physiol.,* 1930, **94**, 41-50.
11. ————. The components of the retinal action potential in mammals and their relation to the discharge in the optic nerve. *J. Physiol.,* 1933, **77**, 207-239.
12. Granit, R., & Hammond, E. L. Comparative studies on the peripheral and central retina: V. *Amer. J. Physiol.,* 1931, **98**, 654-663.
13. Granit, R., & Harper, P. Comparative studies on the peripheral and central retina: II. *Amer. J. Physiol.,* 1930, **95**, 211-228.
14. Kleitman, N., & Piéron, H. Contribution à l'étude des facteurs régissant le taux de sommation des impressions lumineuses de surface inégale. L'hétérogénéité de la rétine. *Année psychol.,* 1928, **29**, 57-91.
15. Koelliker, A. Handbuch der Gewebelehre des Menschen. (Sechste umgearbeitete Auflage.) (Dritter Band von V. v. Ebener.) Leipzig: Engelmann, 1902. Pp. 976.
16. Piper, H. Über die Abhängigkeit des Reizwertes leuchtender Objekte von ihren Flächen bezw. Winkelgrösse. *Zsch. f. Psychol.,* 1903, **32**, 98-112.
17. Ramon y Cajal, S. Die Retina der Wirbelthiere: Untersuchungen mit der Golgi-Cajal Chromsilbermethode und der Ehrlich'schen Methylenblaufärbung. (Trans. by R. Greeff.) Wiesbaden: Bergmann, 1894. Pp. viii+168 (plus diagrams).
18. Riccò, A. *Annalli di Ottal.,* 1877, **6**, fasc. 3, 373. (From Parsons, J. H. *An introduction to the study of colour vision.* Cambridge: Cambridge Univ. Press, 1924. P. 130.)

Clark University
Worcester, Massachusetts

LA SOMMATION SPATIALE DES STIMULI SUBLIMINAIRES DANS LA RÉTINE DE L'OEIL HUMAIN

(Résumé)

On a déterminé les intensités du seuil pour diverses séparations de deux taches de stimulus dans la fovéa et la périphérie de l'oeil humain.

On peut résumer les résultats ainsi:

(1) Dans la périphérie le seuil des deux taches exposées simultanément décroît avec la diminution de la séparation de ces taches (dans les limites de 2½ degrés de l'angle visuel).

(2) Dans la périphérie le seuil des deux taches séparées par une distance de 2½ degrés de l'angle visuel est le même que le seuil de la tache seule qui a la valeur inférieure.

(3) Dans la fovéa le seuil des deux taches exposées simultanément décroît avec une séparation diminuée de ces taches (dans les limites de 10 minutes de l'angle visuel).

(4) Dans la fovéa le seuil des deux taches séparées par une distance de 10 minutes de l'angle visuel est le même que le seuil de la tache seule qui a la valeur inférieure.

On croit que ces résultats montrent l'interaction d'un type sommatif entre des aires stimulées subliminairement. La quantité de cette interaction est une fonction de la distance qui sépare les aires. On offre des raisons pour la conclusion que le lieu de l'effet du seuil visuel se trouve dans les couches synaptiques de la rétine.

BEITEL

DIE RÄUMLICHE SUMMIERUNG NICHT-VERNEHMBARER [SUBBLIMINAL] REIZE IN DER NETZHAUT DES MENSCHLICHEN AUGES

(Referat)

Eben-vernehmbare [threshold] Reizintensitäten sind für verschiedene Entfernungen zweier Reizflecken von einander in der Fovea und an der Peripherie des menschlichen Auges bestimmt worden. Die Befunde lassen sich zusammenfassen wie folgt:

1. An der Peripherie senkt sich, bei simultaner Reizung, die Reizschwelle beider Flecken mit abnehmender Entfernung dieser Flecken von einander (innerhalb der Grenzen eines visuellen Winkels von 2½°).

2. An der Peripherie ist, bei simultaner Reizung, die Reizschwelle für beide Flecke, wenn letztere durch einen visuellen Winkel von 2½° von einander getrennt liegen, die selbe als die desjenigen der beiden Flecke der die niedrigere Schwelle besitzt.

3. In der Fovea senkt sich, bei simultaner Reizung, die Reizschwelle beider Flecke mit abnehmender Entfernung der Flecke von einander (innerhalb der Grenzen eines visuellen Winkels von 10 Minutes) [within 10 minutes of visual angle].

4. In der Fovea ist, bei simultaner Reizung, die Reizschwelle für beide Flecke, wenn letztere durch einen visuellen Winkel von 10 Minuten von einander getrennt liegen, die selbe als die desjenigen der beiden Flecke dem die niedrigere Schwelle eigen ist.

Der Deutung des Verfassers nach, erweisen diese Befunde das Bestehen einer Wechselwirkung [interaction] summierender Art zwischen unterschwellig [subliminally] gereizten Flecken. Die Stärke dieser gegenseitigen Einwirkung ist Funktion der Entfernung der Flecke von einander. Es wird durch Beweisgründe der Schluss unterstützt, dass der geometrische Ort [locus] der Einwirkung der visuellen Reizschwelle [visual threshold effect] in den synaptischen Schichten der Netzhaut liegt.

BEITEL

A STATISTICAL STUDY OF THE RELATIONS OF NEURASTHENIC, DYSPEPTIC, AND ALLERGIC SYMPTOMS*[1]

From the Department of Psychology of The University of Texas

CARLOS KLING

Individual or case examinations of the neurasthenic individual almost invariably yield a certain number of physiological symptoms of disorder in addition to the recognized mental and emotional troubles. It is quite likely, for instance, that the neurasthenic's digestion will be upset and, according to the writer's observations, that he will manifest any number of another class of symptoms which, no matter whether they be produced by reflex nervous mechanisms or directly by a mild anaphylactic reaction to some foreign protein, may be placed under the general category of allergic or hypersensitive disease symptoms. One subject, for instance, who is failing in four-fifths of his courses in college because of easy fatigability, inability to concentrate, day-dreaming, emotional depression, etc., is found to have a variety of intestinal disorders with urticarial rash and also found by skin tests to be specifically sensitive to milk and eggs which are responsible for the intestinal allergy. Both physiological and neurasthenic symptoms disappear with avoidance of these proteins and with the administration of small doses of ephedrine. The subject subsequently passes all his courses with a grade of "B." Another subject with marked emotional instability and depression and most of the other symptoms of neurasthenia reveals upon careful questioning a history of violent migraine attacks and intestinal crises after meals at unaccustomed restaurants. Yet another subject of neurasthenic personality is found by questioning and testing to have an allergic constitution with hay fever, rheumatism, and migraine. Such cases lead one to suspect some causal relations obtaining among these sets of symptoms. Hence the writer was drawn to discover,

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, January 9, 1933.

[1]Part of the routine statistical labor of this project was subsidized by a grant from the Bureau of Research in the Social Sciences.

by the statistical treatment of questionnaires, what degrees of relationship obtain among these three groups of symptoms: the mental and emotional traits of neurasthenia, the physiological symptoms of indigestion, and the physiological symptoms of the allergic diseases.

With this project in view various texts and related sources were consulted for an accepted list of symptoms for each of the groups in question. In the case of the neurasthenic symptoms, only those were selected for the test which were agreed upon as essential, and those others were later culled from the group which did not correlate highly with the group, so that the final items selected to represent neurasthenia are homogeneous to the point of correlating with the rest of the group from .60 to .95, and may therefore be accepted as belonging. Furthermore, inasmuch as the other groups consisted of physiological symptoms, and since it was particularly the relation of what might prove to be the underlying physiological processes to the other symptoms of neurasthenia that interested us, it was well not to reproduce any of those physiological disorders in the neurasthenic group *per se,* else we raise the inter-group correlations spuriously. Consequently, the neurasthenic list was composed of only the mental and emotional traits of neurasthenia with physical symptoms of fatigue only, which symptoms are not duplicated in any of the other groups.

The physiological symptoms were subdivided into groups according to relations obtaining among them. The dyspeptic symptoms were treated statistically in a separate group under the heading of gastro-intestinal symptoms. The total set of allergic symptoms was divided into three groups. Since the isolated symptoms of various allergic diseases would hardly have the same weight as the whole diseases or clinical entities themselves, such as asthma or hay fever, these groups were treated separately. And there are a large number of symptoms of allergy whose occurrence is not general enough to establish them as essential, or which may be reflex or induced; these symptoms were set aside in another group entitled the "allied allergic" for separate correlation with the other groups. This gave us, in all, five major groups: the neurasthenic, the gastric or dyspeptic, the symptomatic allergic, the clinical allergic, and the allied allergic. Besides these groups, manic symptoms from the manic-depressive psychosis were incorporated in the test as a control list, and a number of heterogeneous items of possible interest or value were

lumped together in an unclassified set and treated individually. These seven groups were not separated in the test but were intermixed at random, thus preventing separate group recognition by subjects and forestalling possible prejudices that might affect inter-group correlations.

The test, then, as finally submitted to 808 students in psychology comprised 160 items, grouped as shown in Table 1.

TABLE 1

Neurasthenic	54	items	(27 selected)
Gastro-intestinal	23	"	(12 selected)
Symptomatic allergic	16	"	
Clinical allergic	9	"	
Allied allergic	18	"	
Manic	14	"	
Unclassified	26	"	
Total	160	"	

Characteristics of Allergy

Allergy is defined in general as an altered reactivity of the organism. It was first introduced by Von Pirquet to indicate the remarkable differences in the reactions produced by the first inoculation of an animal with cowpox virus and the second inoculation (3, p. 4). The term refers to the phenomenon of protein sensitization. If a foreign protein is injected parenterally into an animal, no noticeable effects are consequent from this first or sensitizing dose. But if, after a period of incubation of from ten to fourteen days, a second injection of the same protein is given, a number of alarming and often fatal symptoms develop. These are the symptoms of experimental anaphylaxis, and the animal or individual thus sensitized will react to, or be hypersensitive to, this particular protein. This reaction when experimentally produced is given the special term of *anaphylaxis*. Allergy is a more general term which includes anaphylaxis and similar reactions in which the exact origin of sensitization is not known or has not been experimentally produced. The symptoms are varied in kind and degree, and they vary among individuals of species as well as among different species. This same mechanism of sensitization produces the symptoms of a large class of diseases: urticaria, catarrh or rhinitis, hay fever, eczema, catarrhal conjunctivitis, angio-neurotic edema, migraine, serum sickness, asthma, cer-

tain forms of arthritis and rheumatism, epilepsy, and certain obscure abdominal crises of intestinal allergy which give the symptoms of various digestive disorders. These disorders and their symptoms all have a common cause in an altered reactivity of the organism or allergic condition which is usually produced by the direct or indirect inoculation with some foreign protein, but which may in some cases be produced by physical agents, such as heat, cold, sunlight, or exercise (2, pp. 18-19). Sensitivity to a specific antigen or protein is acquired by individuals according to their history in relation to specific causes, but the underlying susceptibility to being sensitized is a condition which is thought to be inherited in Mendelian order as a dominant trait (1). It is the capacity to be sensitized and not any particular disease form that is inherited; a father may have asthma, the son may have eczema, and the grandson may develop epilepsy. Also, the particular form may change in a given individual's life; he may have hay fever as a youth, asthma as an adult, and migraine in old age, or he may have several forms at the same time.

The symptoms resulting from contact with an antigen or sensitizing substance are extremely varied and heterogeneous, depending mainly upon the place of contact and the method of contact. If the reaction is to a food, it will be intestinal and produce intestinal allergy; if to a pollen settled in the nasal mucosa, the reactions will be local hay-fever symptoms; and, if to pollen or animal emanations settled in the lungs, the symptoms will be those of asthma.

Allergic individuals are common among us. The estimates vary from 1 to 15 per cent. The kinds of substances to which one may be sensitive include all the vast forms and kinds of proteins and also a large number of chemicals such as quinine, aspirin, and formalin. A large number of patients are found to be sensitive to purely physical agents such as heat and cold; Duke in particular points out what he calls "reflex-like" symptoms which are mediated by nervous mechanisms and which seem to get attached in a way somewhat analogous to conditioning. Mental or nervous conditions are influential in the appearance of the symptoms and will appear often to be the exciting, if not the sole, cause of attacks.

Procedure

The tests were mimeographed and administered in May, 1931, to all students in psychology and to several classes in educational psychology, the total number amounting to 808. The writer gave

all of the tests himself in order to insure uniformity of presentation. The testing was completed within four days. The students were assured that their scoring would have nothing to do with their grades or their relations with the University, and they were uniformly adjured, for the sake of scientific accuracy, either to score the tests conscientiously and accurately or not at all. They were not required to sign the questionnaire and were told that, if for any reason they did not wish to take the questionnaire, they had only to return a blank copy and no questions would be asked.

The scored tests were separated according to sex and treated separately. Since the groups were intermixed, it was necessary to record the scoring of papers in such a way that they would be more accessible to counting and handling. Accordingly, the scores were transferred to large heavy sheets of finely cross-ruled paper, such that each column might represent an individual and each row a trait. Sheets were procured with 160 rows, exactly the number required for the 160 traits, and with 200 columns. All further statistical counting was done from these sheets. Each paper was checked for errors in this transference.

The first step taken was the correlation of each individual item with the total scores of all 160 items. (In all cases where an individual item was correlated with the total score of a group of items, the coefficient was found by the bi-serial method.) This step was taken to find what traits and what groups of traits contributed most to the final score. The average of the correlations of each item with the total is .46. Since the neurasthenic groups in the correlation of its items with the total yielded the highest average coefficient, this group was taken as the key or master group. This coefficient was .57.

The next step was the purification of the neurasthenic group. This was done by correlating each item of the group with the total of the group. There were 27 out of the 54 items found to correlate above .60 with the group, and these were taken to constitute the final selected neurasthenic group. Then, to see how individual physiological items were related to neurasthenia, all the items were individually correlated with this selected neurasthenic group. Most of the selected neurasthenic items were found to correlate highly with the total score for that group, the coefficients ranging from .70 to .95, with the exception of one, which for some unknown reason fell to .35 in this procedure of refinement. The correlations of individual

items of other groups with the total score of the selected neurasthenic group were found to vary widely; the manic group in particular contributed many negative correlations.

On the basis of these correlations the twelve gastro-intestinal items which had coefficients above .30 with the total score of the selected neurasthenic group were selected. The other physiological or allergic groups, since they were supposedly more heterogeneous, were left intact.

All the groups, with the exception of the unclassified, which had no significance as a group, were next correlated with one another. (All inter-group correlations were obtained by the Pearson product-moment method.) These inter-group correlations are presented in Table 4. In view of the low probable errors, the correlations are statistically significant. There is little doubt that the correlations obtained would have been greater had they not been subject to the many errors of the questionnaire method.

There were very few significant sex differences. The females scored a higher average number of neurasthenic items, the mean for the females being 10.5 and for the males 8.9 items, the P. E.$_{diff.}$ was .306. The males scored consistently a higher number of clinical allergic items, and, although the sex differences for these individual items were not all statistically significant, the fact that all except two of the nine clinical items independently gave differences in the same direction is evidence for a real sex difference.

The emotional depression traits, which showed themselves by their extremely high correlation with the neurasthenic score to be a very integral constituent of neurasthenia, were next selected and correlated as a group with the manic traits. The correlation was .0015, with a P. E. of .026.

The reliability for each group was obtained by correlating the odd with the even items within each group. Unexpectedly high figures were obtained from this measure (Table 2).

TABLE 2

	Raw *r*'s	Corrected *r*'s
Neurasthenic	.7936	.8849
Gastro-intestinal	.5013	.6678
Symptomatic allergic	.5403	.7015
Clinical allergic	.2508	.4010
Allied allergic	.8624	.9261
Manic	.4892	.6569

Finally, the incidence of allergy in its clinical forms and in its symptoms was obtained by tabulating the frequencies of these items. This tabulation indicated an even higher incidence for all forms together as a group than medical estimates have given, the individual items ranging from 2 to 13 per cent of the population.

TABLE 3
INTER-GROUP CORRELATIONS CORRECTED FOR ATTENUATION

	Gastro-intestinal	Symptomatic allergic	Clinical allergic	Allied allergic	Manic
Neurasthenic	.6547	.4489	.3882	.4945	.3348
Gastro-intestinal		.6531	.7429	.6664	.3329
Symptomatic allergic			1.0027	.6857	.2194
Clinical allergic				.6407	.1048
Allied allergic					.2029

TABLE 4
RAW INTER-GROUP CORRELATIONS

Neurasthenic	.5034	.3538	.2117	.4477	.2553
P. E.	.018	.021	.025	.019	.022
Gastro-intestinal		.4471	.3846	.5240	.2205
P. E.		.019	.020	.017	.022
Symptomatic allergic			.5477	.5527	.1490
P. E.			.018	.016	.023
Clinical allergic				.3905	.0538
P. E.				.020	.024
Allied allergic					.1583
P. E.					.023

SUMMARY AND CONCLUSIONS

A list of 160 symptoms of neurasthenia, dyspepsia, and allergy were selected and presented to 808 students in psychology at the University of Texas in the form of a questionnaire, in which the students were asked to indicate what symptoms applied to them. By a preliminary correlation of each item with the total, it was found that the neurasthenic group was the greatest determinant of the final score, and by correlating each neurasthenic item with the total neurasthenic score, 27 items of this group were found to correlate above .60 with the group and were selected to represent neurasthenia. This group was then correlated with a number of other groups and sub-groups of allergy and the symptoms of nervous dyspepsia, and a number of positive correlations of significant magnitude were found. The following are the more important: between neurasthenia and

symptoms of gastro-intestinal disorder, .50; between neurasthenia and allergy (clinical disease forms such as asthma and hay fever), .21; between neurasthenia and two groups of the *symptoms* of allergy, .35 and .45; and between dyspepsia and the above three groups of allergy, .38, .45, and .52, respectively. The correlations of these groups with a control group of manic traits and of a general unclassified number of symptoms yielded low positive, zero, and a few negative relationships. The correlation between the emotional depression traits, which were probably the more fundamental traits of the neurasthenic group, and the manic traits was .0013. The reliability of the test, as measured by correlating odd with even items in each group, was found to be .88 for the neurasthenic group, .67 for the dyspeptic group, .40, .70, and .93 for the three allergic groups, and .66 for the manic group. Very few significant sex differences were discovered. The incidence of hay fever and asthma as forms of allergy is variously estimated by the medical profession, ranging from 1 to 15 per cent. The incidence of the nine clinical forms of allergy listed by the writer indicates even a higher incidence than 15 per cent, if all forms of allergy are thus taken into consideration.

On the basis of the intercorrelations of the symptoms of neurasthenia, dyspepsia, and allergy, one may conclude that certain positive and appreciably significant relationships exist among these groups, but one may conclude nothing concerning the nature of these relations. From the study of individual cases, however, we venture the hypothesis that there are mutual causal relationships among the groups, that many of the physiological symptoms of allergy and dyspepsia have a mental origin in the same personality maladjustments, mental stresses, and emotional strains that may produce neurasthenic symptoms, and that, *vice versa,* a considerable number of neurasthenic symptoms have a physiological origin in obscure and poorly recognized disorders such as allergy and such as individual differences in the physiological capacities and sensitivities of the digestive tract. We are inclined to emphasize the physiological disorders as being the substrata of the neurasthenic constitution; the study of a number of individual cases indicates that lesser degrees and atypical symptom complexes of allergy and of digestive disturbances may be responsible for numerous "mental" disorders of neurasthenia.

Questionnaire

The following is a copy of the test as it was presented to the subjects. The groups of symptoms treated as units are composed of the following items from the test:

Selected neurasthenic group: items numbered 9, 12, 15, 23, 35, 37, 40, 45, 53, 63, 80, 82, 88, 91, 95, 98, 107, 109, 111, 120, 134, 136, 140, 142, 152, 154, 157.

Gastro-intestinal group: items numbered 6, 13, 61, 68, 71, 83, 89, 108, 113, 125, 129, 159.

Symptomatic allergic group: items numbered 4, 5, 14, 29, 34, 38, 54, 55, 58, 62, 64, 77, 100, 101, 122, 123.

Clinical allergic group: items numbered 2, 10, 22, 46, 47, 49, 67, 93, 127.

Allied allergic group: items numbered 1, 18, 21, 24, 30, 39, 42, 43, 51, 65, 72, 76, 81, 121, 133, 135, 141, 143.

Manic group: items numbered 7, 16, 20, 28, 33, 48, 52, 56, 70, 87, 97, 103, 118, 124.

TEST FOR EMOTIONAL TRAITS M F

Instructions

As you know from your own experience, emotional disturbances involve both "mental" and "physical" symptoms. You know also that while some persons are highly emotional, others are emotionally apathetic and cold. All persons, however, show fluctuations in emotional traits.

Please answer the questions frankly and seriously; no one will be able to identify your paper because you will be asked to indicate nothing except your sex. Your instructor will devote at least one class period to a discussion of the results obtained from this test, and you will be interested in the results.

The questions apply only to the present academic year. For example, if you have had *nervous indigestion* at any time during the present academic year, or if you are suffering from it now, encircle the answer "Yes"; if, however, you have had no trace of it during the present year, even if you have had it previously, encircle the answer "No."

The answer "Yes" means that you have or have had enough of the symptom to be noticed by (1) your doctor, (2) your friends or relatives, or (3) by yourself. In other words, "Yes" means that your attention has been called definitely to it. The answer "No" means that, so far as you know, you do not have the symptom now nor have had it during the present academic year.

Please answer each question: it is necessary for us to have both the "Yes" and the "No" answers. Think about each question as you read it, make a definite decision, and then encircle the answer which represents your judgment.

Please consider each item independently of the other items: remember that almost any combination of symptoms may be found in a given individual.

Before you read the questions, encircle either the "M" or the "F" at the upper right-hand corner of this page to indicate your sex. Do this now.

Symptoms

Yes	No	1.	Small pupils.
Yes	No	2.	Urticaria (nettle rash; hives; or "wheels": white or pink swellings on skin.)
Yes	No	3.	Lots of sleep needed in order to feel well; more than most people seem to need.
Yes	No	4.	Hypersecretion of nasal mucous (excessive discharge of secretions from nose).
Yes	No	5.	Sore throat.
Yes	No	6.	Unusual fluctuations in appetite (hunger for food at one time and disgust for food at another).
Yes	No	7.	More or less periodic moods of elation and excitement.
Yes	No	8.	A feeling of internal tension, of being tightly "wound up" on the inside.
Yes	No	9.	Recurrent feelings of mental numbness and inactivity.
Yes	No	10.	Catarrh, coryza, rhinitis, or sinusitis (chronic or constant inflamation or soreness of membranes of nose).
Yes	No	11.	Aversion to routine or regular habits; prefer to eat, sleep, or work when I please.
Yes	No	12.	Dislike for meeting people; usually a wish to avoid them.
Yes	No	13.	Sensations of discomfort, heaviness, and fullness after usual meals.
Yes	No	14.	Desquamatory skin (dry, scaly, and easily chapped).
Yes	No	15.	Feelings of indifference toward everybody and everything.
Yes	No	16.	Tendency to be constantly and physically active; difficult and irksome to be still.
Yes	No	17.	Eye-strain; ache or fatigue from reading.
Yes	No	18.	Extreme sensitivity to dust or smoke of any kind.
Yes	No	19.	Frequent feelings of weariness or fatigue.
Yes	No	20.	High distractibility (attention and thoughts very easily and quickly diverted or changed).
Yes	No	21.	Malnutrition of the scalp (falling hair, or thin hair, or dry and brittle hair, or dandruff, or baldness, etc.).
Yes	No	22.	Hay fever (inflamation of eye and nose membranes with watery secretions, sneezing, etc.).
Yes	No	23.	Moodiness and irritability.

Yes No 24. Voice that is rough, grating, weak, or uneven.
Yes No 25. Blurring of print in reading.
Yes No 26. Fatigue after walking or standing for a short while.
Yes No 27. Extremely sensitive to critical remarks people make about me.
Yes No 28. Rapidly changing and extremely variable emotions.
Yes No 29. Pruritis, or pruritic skin (skin that feels itchy or sticky and uncomfortable).
Yes No 30. Unusually sensitive to noonday light or glare.
Yes No 31. Pyrosis, or "water brash" (bringing up irritating or acid liquids to the mouth in belching or regurgitating).
Yes No 32. Difficulty in speaking under emotion.
Yes No 33. Lack of social inhibitions; quite outspoken, and entirely free in behavior.
Yes No 34. Nasal obstructions; occlusion of sinuses (any cause of stopping of nose passages, such as swelling of membranes with cold, polyps, tumors, deflected septum, or malformed turbinates).
Yes No 35. Periods of depression without apparent cause.
Yes No 36. Bolus hystericus (hard, choking, or tight sensation in the throat, as if the "heart" were in the throat).
Yes No 37. Usually worn out or very tired before the day is over.
Yes No 38. Erythema (reddish, inflamatory, or sore skin).
Yes No 39. Extreme sensitivity of eyes to heat or hot weather.
Yes No 40. Great effort usually required for doing daily work.
Yes No 41. Tendency to be critical of people.
Yes No 42. Extreme sensitivity to winds.
Yes No 43. Pimples (blackheads, swellings, pustules, rashes, acne, or skin infections of any kind).
Yes No 44. Frequent feelings of listlessness, with no desire to do anything.
Yes No 45. Often desires to escape people and be entirely alone.
Yes No 46. Eczema (chronic skin inflamation, scaly, itchy, and eruptive).
Yes No 47. Catarrhal conjunctivitis (reddish, slightly inflamed or swollen membranes of the eyes).
Yes No 48. Unusually accurate memory.
Yes No 49. Angioneurotic edema (puffy, hard swellings arising beneath the skin).
Yes No 50. Auto-criticism (very critical attitude towards self and behavior).
Yes No 51. Ache or pain in eyes.
Yes No 52. Flow of ideas at rapid and high pressure.
Yes No 53. Slowness in making up my mind; difficulty in making ordinary decisions.
Yes No 54. Lacrymation (tendency of eyes to water).

Yes No 55. Itching of the soft palate or pharynx; tendency to "clear the throat."
Yes No 56. Frequent feeling of expansion, confidence, and power.
Yes No 57. Flatulence (gas in stomach or intestines; belching after meals).
Yes No 58. Itch, dryness, or ache of internal membranes of nose.
Yes No 59. Extremely introspective; that is, tendency to study myself constantly.
Yes No 60. Extreme dislike for selling things; for example, would make a poor insurance agent.
Yes No 61. Nervous indigestion (indigestion from worry or moods).
Yes No 62. Sudden sneezing without apparent cause.
Yes No 63. A sense of futility, a feeling that nothing is worth while.
Yes No 64. Itchy, dry, or puffy orbital (eye) membranes.
Yes No 65. Susceptibility to colds (tendency to "take cold," or to have excessive number of colds).
Yes No 66. Absent-mindedness.
Yes No 67. Migraine (spells of sick-headache, or headache with nausea).
Yes No 68. Pain or soreness in abdomen.
Yes No 69. Self-conscious and emotionally upset when required to speak before people.
Yes No 70. Tendency to "put up a good front," "show off," or impress people.
Yes No 71. Heart burn (uncomfortable burning sensations in heart region).
Yes No 72. Spasm of trapezius (pain, stiffness, and soreness in muscles of shoulder and shoulder girdle).
Yes No 73. Often do not get what people say; necessary to ask them to repeat.
Yes No 74. Unpleasant dreams; nightmares.
Yes No 75. Aversion to actions which require initiative and risk.
Yes No 76. Spasm of internal leg muscles (pain, soreness, or stiffness of internal leg muscles).
Yes No 77. Dyspnea (difficulty in breathing or shortness of breath, particularly when recumbent).
Yes No 78. Meniere's syndrome (dizzy spells, with weakness or nausea, after some meals).
Yes No 79. Moods which last sometimes for days.
Yes No 80. Weariness upon arising in the morning.
Yes No 81. Spasm of sterno-cleido-mastoideus (pain, stiffness, or soreness in side and back muscles of neck).
Yes No 82. A feeling of loneliness, even when with people.
Yes No 83. Colitis (inflamation of the large intestine).
Yes No 84. Sometimes made furiously angry over a few trying details.
Yes No 85. Headache.

Yes No 86. Drowsiness or stupor in the daytime, particularly when inactive.
Yes No 87. Rapidly changing interests; flight from one thing to another.
Yes No 88. Difficulty in thinking; forced to grope for ideas.
Yes No 89. Pain, discomfort, or feeling of heaviness or congestion in rectum (lowest length of intestine).
Yes No 90. Spasm of abdominal muscles (pain, stiffness, and soreness in muscles over abdomen).
Yes No 91. Dreading the round of various tasks required from day to day.
Yes No 92. Colic pains (severe griping pains) from eating certain foods.
Yes No 93. Serum sickness (illness from the injection of a serum).
Yes No 94. Nerve aches; nerve pains; neuritis in various parts of the body.
Yes No 95. Frequent disappointment in expectations; things often "go stale."
Yes No 96. Feeling a great relief when the day is over and rest is at hand.
Yes No 97. Talk with great rapidity, force, and excitement.
Yes No 98. Fear that I shall not be prepared for certain demands of the future.
Yes No 99. Hyperchlorhydria (acid stomach).
Yes No 100. Catarrhal deafness (temporary spells of partial deafness, or increased difficulty in hearing).
Yes No 101. Extreme discomfort in certain seasons or localities.
Yes No 102. Pain, discomfort, or weakness without regular meals, or with delay of food.
Yes No 103. Careless, reckless attitude, free of worry or concern.
Yes No 104. Angina pectoris (extreme, sharp pain in heart region, with dizziness, weakness, or fainting).
Yes No 105. Tendency to find sustained physical work extremely fatiguing.
Yes No 106. When sick, tendency to have extreme nausea without relief by vomiting.
Yes No 107. Dread of having to handle new and critical situations.
Yes No 108. Intestinal hypermotility (uneasiness and "growling" of intestines).
Yes No 109. Tendency to be pessimistic and despondent.
Yes No 110. Constipation (less than one *thorough* and *free* excretion of feces daily).
Yes No 111. Feelings of vague uneasiness; distressful feelings that something is wrong.
Yes No 112. Strongly motivated or spurred to effort by praise.
Yes No 113. Bad breath.
Yes No 114. Hypotension (low blood pressure).

Yes No 115. Asthenia (general weakness; easily exhausted; breathlessness and pains from exertion).
Yes No 116. Diarrhea (watery, mush stools).
Yes No 117. Tendency to be peevish and fault-finding in spite of wishing not to be so.
Yes No 118. Tendency to act on impulses to the extent of being somewhat erratic.
Yes No 119. Tachycardia (weak, rapid pulse) after certain foods.
Yes No 120. Difficulty in getting *started* doing things.
Yes No 121. Extreme sensitivity to sound and light; difficulty in going to sleep when there are noises and when light falls on eyes.
Yes No 122. Cough, clearing of throat, or wheezing.
Yes No 123. Severe reactions and illness from insect bites.
Yes No 124. Tendency to be optimistic and expect happiness and success.
Yes No 125. Coated tongue.
Yes No 126. In reading or listening to a lecture, attention lags far behind in spite of conscientious efforts to keep up.
Yes No 127. Asthma (spells of choking, breathlessness, heart fluttering, wheezing, etc.).
Yes No 128. Cold natured (low temperature and easily chilled, especially by cold water).
Yes No 129. Periodic spells of indigestion in spite of care taken to avoid indigestible food.
Yes No 130. Extreme conscientiousness in doing work well.
Yes No 131. Desire to be active and do things, yet constantly falling far short of the goal.
Yes No 132. Food idiosyncrasies (illness from eating certain foods).
Yes No 133. Rheumatism (pain, soreness, and stiffness in muscles and joints).
Yes No 134. Tendency to feel emotionally wretched and miserable.
Yes No 135. Arthritis (inflamation with pain and stiffness in joints).
Yes No 136. Feeling of general inferiority.
Yes No 137. Extremely scrupulous about money matters.
Yes No 138. Often at a loss as to what to talk about with a person.
Yes No 139. Feeling of being unreal, of being depersonalized.
Yes No 140. Tendency to remember humiliating experiences and to feel depressed about them for a long while.
Yes No 141. Irritable bladder (pain in bladder; frequent desire to urinate).
Yes No 142. Tendency to get tired of people quickly and be bored with them.
Yes No 143. Cystitis (inflamation of bladder, urethra, or ureters).
Yes No 144. Dizziness, with objects swimming or misty before the eyes, particularly upon suddenly standing.
Yes No 145. Cold perspiration of hands and soles of feet.
Yes No 146. Self-consciousness in the presence of superiors.

Yes No 147. Insomnia (restlessness and sleeplessness at night).
Yes No 148. Often tormented with emotion I can not express.
Yes No 149. Coldness of hands and feet, very noticeable in cold weather.
Yes No 150. Sexual hypersensitivity (extremely sensitive and easily aroused).
Yes No 151. Syncope (fainting spell).
Yes No 152. Tendency to talk with hesitation and effort.
Yes No 153. Tendency to remain aloof and independent of other people.
Yes No 154. Tendency toward excessive day-dreaming.
Yes No 155. Hemmorhoids (bleeding with defecation or passage of stools).
Yes No 156. Difficulty in articulating words when fatigued or not feeling well.
Yes No 157. Excessive self-consciousness.
Yes No 158. Sleeplessness or nervousness from coffee or tea.
Yes No 159. Indigestion from coffee or tea.
Yes No 160. Tendency towards excessive perspiration.

References

1. Balyeat, R. M. Allergic diseases: their diagnosis and treatment. (3rd ed. of *Hay fever and asthma.*) Philadelphia: Davis, 1930. Pp. 395.
2. Duke, W. W. Allergy, asthma, hay fever, urticaria and allied manifestations of reactions. St. Louis: Mosby, 1925. Pp. 339.
3. Rackemann, F. M. Clinical allergy: particularly asthma and hay fever, mechanism and treatment. New York: Macmillan, 1931. Pp. 617.

Department of Psychology
University of Texas
Austin, Texas

UNE ÉTUDE STATISTIQUE DES RELATIONS DES SYMPTÔMES NÉVROSIQUES, DYSPEPTIQUES, ET ALLERGIQUES

(Résumé)

Puisque les examens cliniques des individus névrosiques montrent bien des symptômes de l'allergie et de la dypepsie, on a fait le projet de découvrir statistiquement les relations entre les conditions allergiques et dyspeptiques et la névrosie. Donc on a préparé une liste de 160 symptômes de ces trois syndrômes et l'a présentée à 800 étudiants en forme de questionnaire. Ensuite on a fait une intercorrélation des groupes de symptômes et des relations positives et signifiantes se sont montrées. Les corrélations suivants, corrigées pour l'atténuation, sont les plus importantes: entre la névrosie et les maladies gastro-intestinales, 0,65; entre les symptômes névrosiques et allergiques, 0,45; et entre les symptômes allergiques et gastro-intestinaux,

0,65. Le coefficient de cohérence de ces groupes de symptômes, mesuré au moyen de corréler les points pairs avec les impairs, a été de 0,88 pour la névrosie, de 0,67 pour la dyspepsie, et de 0,70 pour les symptômes de l'allergie. On suppose que ces relations sont causales et mutuelles. Sans doute les symptômes de l'indigestion et de l'allergie sont produits par les mêmes tensions mentales et les mêmes mauvaises adaptations qui causent la névrosie. L'auteur est disposé à appuyer sur les désordres physiologiques comme base de la constitution névrosique, cependant, et à suggérer que grand nombre de symptômes névrosiques peuvent être causés par des désordres obscurs et mal reconnus tels que l'allergie et les différences individuelles dans les capacités physiologiques et les sensitivités de l'appareil digestif.

KLING

EINE STATISTISCHE UNTERSUCHUNG DER BEZIEHUNGEN ZWISCHEN NEURASTHENISCHEN, DYSPEPTISCHEN, UND ALLERGISCHEN SYMPTOMEN

(Referat)

Da die Prüfung einzelner Fälle neurasthenischer Individuen viele Symptome der Allergie und Dyspepsie aufweist, wurde unternommen, statistisch die Beziehungen der allergischen und dyspeptischen Zuständen zur Neurasthenie zu entdecken. Es wurde daher eine Liste von 160 Symptomen dieser drei zusammentreffenden Symptomgruppen vorbereitet und 800 Studenten in der Form eines Fragebogens vorgelegt. Die Symptomgruppen wurden dann interkorreliert, und man fand, dass die positiv und bedeutsame Beziehungen aufweisen. Die folgenden, auf Verminderung korrigierten Korrelationen, sind die wichtigsten: zwischen neurasthenischen und Magen- und Darmstörungen, 0,65; zwischen neurasthenischen und allergischen Symptomen, 0,45; und zwischen allergischen und Magen- und Darmsymptomen 0,65. Der Zuverlässigkeitskoeffizient dieser Symptomgruppen, gemessen durch das Korrelieren der geraden und ungeraden Angaben, war 0,88 für Neurasthenie, 0,67 für Dyspepsie und 0,70 für die Symptomen der Allergie. Man nimmt an, dass diese Beziehungen kausal und gegenseitig sind. Symptome der Verdauungsschwäche und der Allergie werden zweifellos durch dieselben geistigen Spannungen und Fehlanpassungen hervorgerufen, die die Neurasthenie verursachen. Der Autor neigt aber die physiologischen Störungen als Substrate der neurasthenischen Konsitution zu betonen und die Vermutung aufzustellen, dass eine beträchtliche Zahl von neurasthenischen Symptomen durch dunkel und schlecht erkannte Störungen, wie die der Allergie und individuelle Unterschiede in den physiologischen Fähigkeiten und der Empfindlichkeit der Darmkanäle verursacht werden.

KLING

INDIVIDUAL DIFFERENCES IN ABILITY TO ALTERNATE ACTIVITIES*[1]

From the Psychological Laboratory of the University of Chicago

CLARA BURRI

PROBLEM

The problem with which this study is concerned may be formulated as follows: The primary purpose of this experimental and statistical investigation is to find out whether people differ consistently in their ability to do varied as against constant tasks, and to what extent this ability to alternate is a function of the tasks and conditions. This question we attempted to answer by testing to see how the ability is influenced by such variables of the alternating situations as the kind of tasks; the condition of alternation, that is, whether alternation occurs between similar or dissimilar tasks; and the frequency of alternations. This last aspect has been merely touched upon in this study. A second objective of the study is to throw more light on the problem of whether individual differences in ability to alternate between tasks may be wholly due to the degree of perseverative activity, or whether some additional factors are present, and, if so, what kind. A third objective of the study is to see whether, on the average, continuous or alternating activity is more effective, and whether the average results remain constant from task to task.

METHOD AND PROCEDURE

Material and Subjects. The general method employed is that of individual experimentation in which simple tasks were performed singly and in alternation. Fifty-one undergraduate and graduate students from the University of Chicago were given four pairs of tests which had been selected to represent various degrees in similarity

*Recommended by A. G. Bills, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, June 27, 1933.

[1]A part of the dissertation submitted to the Faculty of the Division of the Biological Sciences in Candidacy for the Degree of Doctor of Philosophy in the Department of Psychology, University of Chicago.

of material and various differences in kinds of tasks. We also attempted to choose as nearly as possible, without interfering with the specific problem involved in this study, the same activities which have been used before in experimental studies of alternating activity. In the first pair of tests, the subjects had to work either continuously on simple addition or subtraction problems of two-place numbers or to shift from addition to subtraction every minute or every ten seconds. For the second test-pair the same procedure was employed, only this time the task was to write either adjectives or verbs as associative responses to a list of given nouns. These two sets of tests were selected to represent predominantly mental tasks, while the third set of tests was predominantly motor. It consisted in drawing triangles inside circles, continuously putting the point up or down, or alternating between the two methods. Finally, in the fourth test-pair the subjects were required to form S's or closed figures (✡) from groups of six dots which had been placed together without definite arrangement. Those tests required perception of form or configuration. When the two figures are formed in alternation it not only indicates the ease with which a given individual can perceive one or another form from a group of patternless dots, but also his ability to change motor set.

Conditions. Three main conditions of work were used:

A. Continuous work
B. Alternation between two tasks, every minute
C. Alternation between two tasks, every ten seconds

The work material was combined into groups of similar and dissimilar tasks. The alternation activity of similar tasks was:

1. Additions—subtractions
2. Verbs to nouns—adjectives to nouns
3. Drawing triangles, point up—drawing triangles, point down
4. Forming closed figures—forming S's

The dissimilar tasks consisted of:

1. Subtractions—verbs to nouns
2. Additions—forming closed figures
3. Adjectives—drawing triangles, point down
4. Drawing triangles, point up—forming S's

Procedure. Every subject worked for eight minutes at a given

task either continuously on one task for the whole period or alternating between two tasks. During continuous work he had to draw a line after the last element finished whenever he heard the signal "check" given by the experimenter. This checking signal was given every minute, thus allowing an analysis of the amount of work performed during small intervals. When alternating work was done, a changing signal was given either every minute, or every ten seconds, depending upon the condition. When changing every minute the subject also made a check ten seconds after every change. This procedure gives a check on the immediate effect of changing.

In order to counterbalance practice effects every subject went through every test twice, first in normal and then in reversed order. As the final score the average amount of work from the two performances was used. The conditions were carried out in the following order:

A. Continuous work
1. Adding
2. Adjectives to nouns
3. Triangles, point up
4. Forming closed figures from groups of six dots
5. Subtracting
6. Verbs to nouns
7. Triangles, point down
8. Forming S's from groups of six dots

B. Alternating every minute
1. Similar tasks
   a. Subtracting—adding
   b. Adjectives to nouns—verbs to nouns
   c. Triangles, point up—triangles, point down
   d. Forming closed figures—forming S's
2. Dissimilar Tasks
   a. Subtracting—verbs to nouns
   b. Adding—forming closed figures
   c. Adjectives to nouns—triangles, point down
   d. Triangles, point up—forming S's from dots

C. Alternating every ten seconds

For this part of the investigation the same tests and order of procedure as for alternating every minute were used. Only six subjects participated in this condition.

After a subject finished this order, every test was repeated; but this time he started with the condition and tests finished at his last sitting and ended with continuous adding, the task which before he had performed first.

The various problems were presented on sheets of paper on which the subjects also wrote the answers with a soft pencil. In the construction of arithmetic and word tests care was taken that a given problem or work was not repeated. When changing between similar tasks, such as adding and subtracting, the work was done on one sheet, simply by changing from the process of adding to that of subtracting, or vice versa, when the signal was given to change. But when working on alternation of dissimilar material the subjects wrote on two sheets, side by side, one containing, let us say, subtractions; and the other, groups of dots. Under this condition the subjects worked first for a minute on one sheet and then for a minute on the other, thus going back and forth from sheet to sheet whenever a change of task was demanded. From the point of view of an ideal experimental procedure, the use of these two methods is not desirable, but, since in alternating from sheet to sheet some trial subjects took about the same length of time as they did to start working again on the same sheet, it was decided to keep one sheet wherever it was possible. By so doing, the scoring and keeping-track of the combinations was greatly simplified.

Each task was continued for eight minutes, during which time a subject worked either alternatingly or continuously. The score for one working period was the number of additions or figures or verbs, etc., produced in four minutes, because, in alternating, half of the working period is filled with doing one task and the other half with another similar or dissimilar task. Continuous work was kept up for eight minutes rather than for only four in order to equalize fatigue effects. The score from eight minutes of work was then simply divided by two to make it comparable with alternation scores. No use of errors was made.

Treatment of Data. In order to obtain comparable measures of efficiency in doing alternate versus continuous work, a measure which is not influenced by the total amount of work, the ratios between the results from the two ways of working were computed. Taking the results from the continuous tasks as a standard, a subject could obtain a plus or minus score for alternating work, depending

upon whether he accomplished more or less when alternating between two activities than when working continuously on one task. For example, if a subject has a score of +8 per cent for adding, that means that he did 8 per cent more work when he alternated adding with subtracting than when he only added. A minus score, let us say, —10 per cent for adding, means that, when alternating, 10 per cent less work was performed by the subject than when working continuously. Every subject received such a ratio for every activity, first, for every task when alternated with its similar task, such as adding—subtracting; adjectives—verbs; triangles, point up—triangles, point down; S—closed figure; and, second, for every task when alternated with a dissimilar task, such as subtracting—verbs; adding—closed figure; adjectives—triangles, point down; and S—triangles, point up. Thus, there are sixteen ratios for every subject. In order to see whether some consistent trend is present for all subjects and whether a general factor of alternating ability may be found for all the tests, the intercorrelations and tetrad differences were computed. A multiple-factor analysis was also made.

Results—For the Group as a Whole

Changing between Similar Tasks versus Continuous Performance. Did the subjects taken as a group work better when changing every minute from one to another similar task, or when working continuously on one task? Table 1 shows an answer to this question. It gives the average ratios from the pool of the eight tests. Thirty-eight of the 51 subjects obtained a positive average ratio from the eight tests, which means that, on the average, they worked better when alternating between two tasks than when working continuously on one task, while the remaining 13 subjects obtained a negative average ratio from the eight tasks, indicating that they did better on constant work. Considering the group as a whole, we obtain an average of 2.80 per cent in favor of alternation. This difference of not quite 3 per cent is not statistically significant, but since 38 of the 51 subjects are better at alternating than at continuous work it indicates a definite trend.

Although the pooled results of the group indicate only little advantage of alternation over continuous work, it is possible that if we study the results from the various tests separately this tendency may come out more clearly. Perhaps with some tasks one can

TABLE 1

AVERAGE OF RATIOS COMPARING ALTERNATING WITH CONTINUOUS PERFORMANCE ON EIGHT TESTS USING SIMILAR TASKS

Subjects	Average improvement alternating	P.E. of average
38 S's better alternating	5.82	±.468
13 S's better continuous	6.00	±.675
All subjects	2.80	±.769

achieve more when one is alternating between two similar activities while with another activity it may be better to work continuously in order to obtain maximal efficiency. Some light is thrown on this problem by Table 2.

TABLE 2

THE RATIOS BETWEEN PERFORMANCE ON CONTINUOUS AND ALTERNATING SIMILAR TASKS

Tasks (and tasks with which alternated)	Average ratios	P.E. of average	Sigma P.E.
Adding—(subtracting)	—1.25	± .97	10.30± .68
Subtracting—(adding)	—3.35	± .97	10.22± .68
Adjectives—(verbs)	1.80	±1.09	11.05± .77
Verbs—(adjectives)	2.05	±1.37	13.50± .87
Triangles, point up—(triangles, point down)	9.00	±1.56	16.50±1.09
Triangles, point down—(triangles, point up)	3.75	±1.09	11.50± .76
Closed figure—(S)	5.05	±1.10	12.00± .80
S—(closed figure)	6.20	±1.04	12.50± .82

Table 2 gives the average differences for each test separately. For six out of the eight tasks the group did better, on the average, when alternating. For addition and subtraction there is, however, a small difference in favor of continuous work. Although the average ratios of —1.25 per cent for addition and —3.35 per cent for subtraction are not statistically significant, it is nevertheless interesting to note that these results agree with those obtained and reported by Dashiell (*2a*). He too found that his subjects added better when they were permitted to complete one hundred problems continuously than when they were required to mix the additions with subtractions, divisions, and multiplications. It is possible that, when addition is alternated with subtraction, interference of specific responses occurs. This suggests the necessity of special explanatory factors adapted to particular situations. The group averages for the different tasks suggest that

perhaps the kind of task is a factor in determining whether a group of subjects is more efficient when alternating between two similar activities or when working continuously on one. There is a definite difference in degree of advantage from alternating work for the tests involving verbal material and those for motor activity, the latter being consistently larger.

TABLE 3

AVERAGE OF RATIOS COMPARING ALTERNATING WITH CONTINUOUS PERFORMANCE ON EIGHT TESTS USING DISSIMILAR TASKS

Subjects	Average improvement alternating	P.E. of average
47 S's better alternating	+10.65	± .63
4 S's better continuous	— 3.28	±1.11
All subjects	+ 9.55	±1.28

Changing between Dissimilar Tasks versus Continuous Performance. If our subjects alternate between two dissimilar tasks, such as addition and drawing a figure, we find that most subjects perform more work when they can alternate between tasks at intervals of one minute than when they have to work on one task for eight minutes at a stretch. On an average, 9.55 per cent more work was completed during a four-minute working period when the subjects alternated.

This superiority of alternating over continuous work may be seen even more clearly from Table 4, where the group averages are reported for each task separately. As far as these tasks are concerned, varied activity consistently brings about greater efficiency than does continuous activity.

In general, these group results agree with the findings of Wyatt (5, 6), Chapman (2), Robinson and Bills (4), Bills and McTeer (1), and others. These authors all found that people tend to work better when alternating, but that this superiority of alternation varies from task to task and from situation to situation, depending upon various factors. Those which are shown to have the greatest influence on the effect of alternation versus constant activity are: (1) the kind of the tasks to be alternated, whether arithmetic, or adjectives, or triangles; and (2) the similarity of the tasks, whether alternation occurred between two kinds of arithmetic, or between arithmetic and something else.

TABLE 4

THE RATIOS BETWEEN PERFORMANCE ON CONTINUOUS AND ALTERNATING DISSIMILAR TASKS

Tasks (and tasks with which alternated)	Average ratios	P.E. of average	Sigma	P.E
Adding—(closed figure)	10.54	±1.52	14.00	± .93
Subtracting—(verbs)	8.97	± .94	10.00	± .68
Adjectives—(triangles, point down)	11.42	±1.11	11.80	± .78
Verbs—(subtracting)	8.85	±1.45	15.40	±1.03
Triangles, point up—(S)	10.25	±1.56	16.50	±1.10
Triangles, point down—(adjectives)	3.77	±1.41	14.95	±1.00
Closed figure—(adding)	7.40	± .96	10.45	± .69
S—(triangles, point up).	17.40	±1.51	16.05	± .77

Individual Differences in Ability to Alternate. In this study interest is directed mainly to the problem of individual differences. Instead of studying merely group results, an attempt was made to find out whether a difference exists in ability to do varied as against constant tasks. If so, is this difference a consistent one, or does the ability to do alternating work vary from condition to condition? If it does vary, what are some of the factors that influence the ability? An attempt is made to answer these questions in the following part of this report.

Types of alternators. First let us see whether subjects can be arranged according to type. We have already shown that we did not find 100-per-cent consistency in our subjects' ability to alternate. With the exception of one subject the ability to alternate varied from task to task. Sometimes a subject worked better alternately between two tasks, at others continuously at one. However, when we calculated some sort of average performance score for each subject we found that the subjects showed a trend either for being better at alternation or at constant work.

Since no complete consistency in a person's ability to alternate could be found, the problem next arises of deciding on what basis to divide the group of subjects. This may be done in several ways, and the results will vary according to the criterion of classification. For the sake of convenience, let us classify a subject as a good or a poor alternator according to whether he did better one way or the other in more than half of the eight tests, and also according to whether his average score or ratio from the eight tests is negative or positive.

If such a classification is carried out regardless of the amount of difference between alternating and constant work, and regardless of the consistency in tendency from test to test, it is found that 37 subjects did better alternating regardless of the condition, 10 subjects were good alternators when the change occurred between dissimilar tasks, but not between similar tasks; while only 3 out of the 51 subjects did better on constant work regardless of the condition. One subject worked better when alternating similar tasks, while he did less well when alternating dissimilar tasks. However, in this last case the difference between the two conditions is so small that it can be neglected. Table 5 gives the average of ratios comparing continuous and alternating similar and dissimilar tasks for each of these groups, the poor and the good alternators.

TABLE 5

AVERAGE OF RATIOS COMPARING CONTINUOUS WITH ALTERNATING SIMILAR AND DISSIMILAR TASKS, FOR POOR AND GOOD ALTERNATORS

"Type"	Alternating	
	Similar tasks average ratios	Dissimilar tasks average ratios
1. Good alternators (37)	+5.9	+11.6
2. Good alternators in condition two, poor in condition one (10)	—5.5	+ 7.2
3. Poor alternators (3)	—6.1	— 4.2
4. Good alternator in condition one, poor in condition two (1)	+2.1	— 0.5

Although we showed in the above paragraph that some people tend to be good or poor alternators, the average ratios are insignificantly small. Furthermore, we must remember that this classification is only an approximate one. In order to show just how inconsistent these people were in their ability to alternate, we tabulated the number of times each task was performed better during alternation and the number of times each task was performed better during constant work. These figures are shown in Table 6.

The number of times each test came out in favor of alternation is much more frequent and consistent for continuous versus alternating dissimilar tasks than for continuous versus alternating similar tasks. Additions and subtractions in the test for alternation of similar tasks are the only ones which turned out more than half the time in favor

of constant work. If we remember, the group averages, too, show this same tendency.

TABLE 6

THE FREQUENCY WITH WHICH EACH TASK WAS PERFORMED BETTER ALTERNATING THAN CONTINUOUSLY, AND VICE VERSA

Task	Continuous *vs.* alternating similar tasks and dissimilar tasks			
	Continuous alternating	Alternating continuous	Continuous alternating	Alternating continuous
Adding	29	22	10	41
Subtracting	35	22	9	42
Adjectives	20	31	11	40
Verbs	25	26	14	37
Triangles, point up	15	36	14	37
Triangles, point down	18	33	21	30
Closed figure	14	37	10	41
S	16	35	8	43

The effect of the kind of task. In the previous analysis of the group results we found that in general the ability to alternate differed from task to task. We remember that by varying between two similar tasks the subjects proved to be slightly more efficient when performing addition and subtraction as constant rather than as alternating tasks. Although all the other tests gave results in favor of alternation, the degree of this advantage varied for different tests, and for the two conditions, similar and dissimilar alternated tasks. If the ease of shifting varies in general from task to task, how then does the task influence an individual's ability to alternate? Will different subjects who are more efficient on the average when doing varied as against constant work show an equal degree of efficiency for all the tests? Or will the degree of their ability to change vary from task to task? Or may it even be absent for some tasks but present for others? These same questions arise for poor alternators, subjects who are more efficient at constant work as compared with varied work. In order to determine the degree of consistency from task to task the intercorrelations for the various tests were computed. They are presented in Tables 7 and 8.

Most of the correlations are positive, even though rather small. For continuous versus alternating similar tasks the coefficients range from —.08 to +.64, with an average coefficient of .21. Only 6 of the 28 correlations are four or more times their P.E.'s. The table of intercorrelations for continuous versus alternating dissimilar tasks

shows in the main the same tendency. Here we have a range of coefficients from —.11 to +.72, with an average correlation of .17. Again only 6 of the 28 coefficients are statistically significant, being four or more times their P.E.'s. These intercorrelations furnish little evidence of a general ability to alternate tasks. According to these results, if a person does better when alternating on one test it cannot be assumed that he will also do better when alternating on another test.

TABLE 7

INTERCORRELATIONS FOR THE RATIOS BETWEEN CONTINUOUS AND ALTERNATING SIMILAR TASKS

	1	2	3	4	5	6	7	8
1. Adding	—							
2. Subtracting	.44	—						
3. Adjectives	.29	.03	—					
4. Verbs	.24	.16	.37	—				
5. Triangles, point up	.07	.22	.03	.03	—			
6. Triangles, point down	.19	.21	.02	.04	.64	—		
7. Closed figure	.09	.22	.10	.11	.28	.35	—	
8. S	—.08	.18	.15	.12	.34	.26	.55	—

TABLE 8

INTERCORRELATIONS FOR THE RATIOS BETWEEN CONTINUOUS AND ALTERNATING DISSIMILAR TASKS

	1	2	3	4	5	6	7	8
1. Adding	—							
2. Subtracting	.14	—						
3. Adjectives	.09	—.09	—					
4. Verbs	—.10	—.11	—.06	—				
5. Triangles, point up	.21	.23	—.07	.27	—			
6. Triangles, point down	.02	.29	.05	.01	.42	—		
7. Closed figure	.15	.30	.09	.09	.72	.25	—	
8. S	.02	.32	.27	.14	.63	.21	.42	—

By inspection it may be noticed that the interrelations from predominantly "motor" tasks show some consistency; they also give the highest coefficients. Further, there is also some slight consistency in relationship between these tests and subtraction, but not between them and addition or verbs or adjectives. Another group of correlations which might indicate some degree of relationship, but to a slighter degree than the others, are those from the predominantly "mental" tasks, from the groups of alternating similar activities. A

comparison of the two sets of intercorrelations from alternating similar versus constant activity and from alternating dissimilar versus constant activity indicates little consistency of relationship for the two conditions of alternation. Again, a comparison of the intercorrelations from the predominantly "motor" and "mental" tasks suggests the idea that different factors are operating in these two sets of correlations.

The relatively high correlation for the "motor" tasks in the case of dissimilar alternation may be due to similarity of movements in these activities. For example, the highest correlation, .72, is between drawing triangles (↓Δ) in the direction as indicated by the arrow, and drawing figures (↓⇧) where we find the same general direction of movements. The same similarity of movement was used in drawing triangles (↓Δ) and S's (↓S), and these tests give a correlation of .63. According to these results it seems that the similarity of the movements in the response may be an important factor in determining an individual's ability to change from task to task.

Is there a general factor of alternating ability? From the procedure employed above, i.e., merely inspecting the tables of intercorrelations, it cannot be said whether or not one general factor and some specific factors are sufficient to account for these intercorrelations. In order to detect whether or not a common factor is present, the Spearman two-factor analysis was made. The tetrad differences of the intercorrelations for continuous versus alternating similar tasks may be summarized as follows:

The average deviation of tetrads from zero = .0348
The theoretical P.E. =±.0395
The largest tetrad = .2669±.069
The largest tetrad is 3.8 times its P.E.

For continuous versus alternating dissimilar tasks the results are:

The average deviation of tetrads from zero =—.0470
The theoretical P.E. =±.0377
The largest tetrad =—.2654±.067
The largest tetrad is 3.95 times its P.E.

According to Spearman and Holzinger, if the tetrad differences are to be explained as chance variations from zero, the average tetrad difference should be the same size as the theoretical average probable

error.[2] Another way of testing whether the two-factor criterion is satisfied is by seeing whether half of the tetrad differences are larger and half of them smaller than the P.E. Still another summary test is to take the largest tetrad difference and compute its P.E. If this tetrad difference is less than two and one-half times its P.E., it may be considered as insignificant. If, however, this largest tetrad difference is significantly large, five or more times the P.E., or at least if it is too large to assure that the two-factor analysis holds, then some group factor or factors may be present.

Accordingly, our results may be interpreted as satisfying the two-factor theory. At least we may say that most of the tetrad differences would vanish, a fact which indicates that most of them are caused by chance variations from zero. However, at the extremes of the distributions of the tetrad differences there are a few which are fairly large, and which suggest, therefore, the possibility of some additional factors. Since as a whole the dispersions of the tetrad differences which were obtained compare very closely with those dispersions which would be expected if the differences varied from their theoretical zero by chance factor only, it may be concluded temporarily that one factor is sufficient to account for the observed intercorrelations. Since the variables were all tests of alternation, let us call this common factor *A,* or alternating factor.

Our next problem is to determine the relation of this common factor to any of our variables. The correlations between any one of the variables and the common factor may be determined by finding the interrelations of any three of the variables, by the formula $r_{kg}=\frac{rak \quad rbk}{rab}$. By this method we estimated the correlation between a given test for the condition alternation of similar tasks and the common factor. The results are:

	r_{kg}
Addition	.5203
Subtraction	.3797
Adjectives	.2656
Verbs	.4170
Triangles, point up	.4903
Triangles, point down,	.4806
Closed figure	.5631
S	.6950

[2]See formula 16 in Holzinger's (3) monograph on the tetrad-difference criterion.

The relationship of this common factor to the different tests varies from test to test. For subtraction and adjectives it is rather low; in fact, it is not present to any great extent anywhere except in the S-forming test.

The significance of this factor loading was estimated by the formula $\Sigma rkg^2/n$. The index of significance for the common factor is .2439, which indicates that, although a common factor is present, it is not very important in accounting for the intercorrelations for the condition alternating similar versus constant tasks. Presumably for the majority of these tests the specific factors are equally as important as, or even more important than, the common factor in accounting for the intercorrelations.

Because of these findings the question arose whether these tests could not be better described by several factors than by a single factor. Although Spearman's criterion is satisfied, it is entirely possible that some of these tests have something else in common. Therefore we proceeded to try a multiple-factor analysis.

A multiple-factor analysis. The technique for this multiple-factor analysis is the "simplified method" of Professor Thurstone.[3] Since the method is intricate and lengthy, a description of its application would be out of place here. Hence, we shall merely report the results.

By investigating our two sets of tests with a multiple-factor pattern it was discovered that each set has one conspicuous common factor running through all of the tests and that in addition each set had three other factors of more or less importance. The four factor loadings for the eight variables for the two sets of tests are as follows:

1. Continuous versus Alternating Similar Tasks

	k_1	k_2	k_3	k_4
Adding	.4278	.5353	—.2525	.1498
Subtracting	.4838	.2119	—.2725	.3728
Adjectives	.3436	.3803	.3782	—.1751
Verbs	.3820	.3803	.2857	—.1520
Triangles, point up	.5882	—.2913	—.2005	—.1912
Triangles, point down	.5984	—.2793	—.2835	—.1166
Closed figure	.5730	—.3163	.2262	.3048
S	.5271	—.3736	.3493	.2114
Index of significance ($\Sigma k^2/n$) =	.2491	.1252	.0845	.0502

[3]The procedure for the application of this technique was obtained from a lecture on this subject, which will soon appear in print.

2. Continuous versus Alternating Dissimilar Tasks

	k_1	k_2	k_3	k_4
Adding	.1600	.0938	—.0801	—.1008
Subtracting	.4516	.2184	—.5140	.1286
Adjectives	—.0029	.7940	—.5252	.2480
Verbs	.1743	—.3055	.3725	.1994
Triangles, point up	.9349	—.0095	.3005	.1809
Triangles, point down	.4430	—.0386	—.1768	—.0237
Closed figure	.7316	.1232	.0365	.2019
S	.6002	.1892	.0993	.4642
Index of significance $(\Sigma k^2/n)$ =	.2782	.1052	.1020	.0543

These results indicate that the ability to alternate is considerably more complex than has hitherto been assumed. The nature of the factor loadings for condition alternating similar versus constant activity suggests that the first factor may be ability to alternate muscular set, and the second one possibly represents flexibility in symbolic processes. Its presence increases ability to alternate addition, subtraction, adjectives, and verbs, while it diminishes ability to alternate the more muscular activities. The third and fourth factors are of minor significance and have no clear interpretation. They do, however, indicate the presence of other variables which help determine ability to alternate.

The factor pattern from alternating dissimilar versus constant tasks shows the same general tendency as does that for alternating similar versus constant activities. Here, too, we find that there is one big factor which is common to all the tests, and three factors of less importance, two of which are of equal importance but less than the first, while the third one seems to be of only minor importance. It is somewhat difficult to interpret these factors. However, we may say that the first factor loading is again largest for the predominantly motor tasks.

Whatever we may decide regarding the nature of some of these details, we may at least draw a few conclusions about the findings as a whole, namely: Ability to alternate is very complex; there are at least several factors which determine whether an individual will work better alternatingly or continuously; for these results, four factors for each set of variables have been isolated. The most important of these factors, and the only one which has a clear rational

interpretation, seems to be in the nature of ability to change motor set. In the case of alternating similar tasks, there appears to be a second factor parallel to this first which may represent flexibility on the mental side.

The effect of various time intervals on ability to alternate. In order to see whether different lengths of time intervals between changes produce different effects on ability to alternate from task to task, six students worked the eight tests under two different conditions, at one time changing every minute and at another time every ten seconds. In Table 9 the average ratios between continuous activity and alternating between two similar and two dissimilar tasks every minute or every ten seconds are given for each subject and for the group.

TABLE 9

AVERAGE RATIOS BETWEEN CONTINUOUS AND ALTERNATING SIMILAR AND DISSIMILAR TASKS WITH CHANGE EVERY MINUTE AND EVERY TEN SECONDS

	Alternating			
	Every minute		Every ten seconds	
Subjects	Similar tasks ratio	Dissimilar tasks ratio	Similar tasks ratio	Dissimilar tasks ratio
1	—4.6	8.7	2.1	14.6
2	3.3	16.0	10.3	12.5
3	5.2	12.5	5.0	3.3
4	8.7	11.8	20.3	17.5
5	6.7	2.2	4.8	4.8
6	3.4	11.6	17.0	25.1
Averages	3.8	10.5	7.2	12.9

The general tendency of the results is the same when the change was made every ten seconds as when it was made every minute. The group averages are in favor of alternating for both conditions and both time intervals. Again, we find that the average difference between constant and variable work is greater for alternating dissimilar tasks than for similar tasks. This tendency exists when change occurred each minute and when it took place every ten seconds. When the subjects alternated every ten seconds they accomplished more than when they changed every minute. This is probably because they put more effort into their work; for when they alternated every ten seconds small enough work segments were produced so that they were able to keep track of their own accom-

plishment, and they tried to increase their speed throughout a work period. The reports indicated clearly that some of the subjects tried to do just a little more than usual. This special effort, rather than the time interval as such, may have caused the greater increase in efficiency for the condition alternating every ten seconds.

General Summary

In this investigation we have attempted to throw some light on the problem of ability to alternate activities in respect to these three considerations: first, whether individuals differ consistently in their ability to do varied as against constant work, and hence whether or not they may be considered either good or poor alternators; secondly, whether individual differences in ability to alternate may be considered as being due entirely to the degree of perseverative activity as certain investigators have assumed, or whether other factors are influential; and, thirdly, whether, on the average, continuous or alternating work is more effective.

1. In regard to our first problem we found that although people differ greatly in their ability to alternate activities they cannot be considered as good or poor alternators in a general sense. They show little consistency in their ability to do varied as against constant tasks. A given individual may be a good alternator in one situation and poor in another one. Ability to alternate was found to depend upon the type of the tasks alternated, whether arithmetic, association, or drawing triangles or figures, and upon whether the shift occurred between similar or between dissimilar tasks.

Even if we adopt a very loose criterion of "type," there seemed to be little tendency for a subject to be either a good or a poor alternator. When we tried to classify our subjects into poor and good alternating types on the basis of whether or not they did better alternating on at least four out of the eight tests and whether their final average performance was in favor of alternating or continuous work, we found that in alternating similar tasks 38 out of our 51 subjects worked better alternatingly and 13 worked better continuously, while for alternating dissimilar tasks 47 worked better, on the average, at alternation and only 4 at constant activity. Even according to this loose classification only four subjects can be regarded as poor alternators within the conditions of our experiment.

2. According to our results the problem of individual difference

in alternating activities is much more complex than has hitherto been assumed by those investigators who have attributed it wholly to differences in degree of perseverative activity. Our experimental findings agree with the conclusions reached in our analytical study,[4] in which we showed that, although "perseveration" may contribute to whether or not an individual is a good alternator, it is by no means the only, nor even the principal, causal factor.

By the application of the method of multiple-factor analysis to our data we isolated four factors for each of our two sets of alternating conditions (similar and dissimilar tasks). Each set has one conspicuous common factor and three other factors of more or less importance. The nature of the factor loadings suggests that the leading ones may be ability to change a general motor set and flexibility in alternating symbolic material.

3. On the average, our subjects worked better at alternating than at continuous tasks in the case of both similar and dissimilar tasks. However, the superiority of alternation depends in part upon the particular tasks. For example, for arithmetic tests our subjects gave an average performance in favor of continuous activity. The degree of superiority of alternation was influenced by the degree of similarity of the alternated tasks. If alternation took place between dissimilar tasks, the group, on the average, did 9.55 per cent better when alternating than when working continuously, while for alternating similar tasks it was only 2.80 per cent better.

Varying the frequency of alternation from one minute to ten seconds produces little difference in results for alternation versus constant activity. The advantage of alternating work is even greater when alternation takes place every ten seconds than when it takes place every minute.

References

1. Bills, A. G., & McTeer, W. Transfer of fatigue and identical elements. *J. Exper. Psychol.*, 1932, **15**, 23-36.

2. Chapman, J. G. The effect of rapid change of work on the rate of performance. *J. Exper. Psychol.*, 1917, **3**, 165-170.

2*a*. Dashiell, J. F. Fundamentals of psychology. New York: Macmillan, 1920. (See esp. Chap. 10.)

[4]A critical review of the "present status of individual differences in alternating activities" to be published in the *Psychological Bulletin.*

3. Holzinger, H. S. Statistical résumé of the Spearman two-factor theory. Chicago: Univ. Chicago Press, 1930. Pp. 43.
4. Robinson, E. S., & Bills, A. G. Two factors in the work decrement. *J. Exper. Psychol.*, 1926, **9**, 415-443.
5. Wyatt, S. The effect of change in activity. (*Indus. Fatigue Res. Bd. Rep.*, No. 26.) London: H. M. Stationery Office, 1924. Pp. 13-37.
6. Wyatt, S., & Fraser, J. A. The comparative effects of variety and uniformity in work. (*Indus. Fatigue Res. Bd. Rep.*, No. 52.) London: H. M. Stationery Office, 1928. Pp. iii+30.

c/o Ernest Bauman
North Adams, Michigan

LES DIFFÉRENCES INDIVIDUELLES DANS LA CAPACITÉ D'ALTERNER LES ACTIVITÉS

(Résumé)

Cinquante-et-un étudiants universitaires ont fait un travail continu d'addition, de soustraction, de placer un verbe ou un adjectif avec un substantif, de dessiner des triangles de forme ordinaire ou renversée, de former deux sortes de figures en employant des groupes de six points, et ensuite ils ont alterné entre deux de ces activités, telles que l'addition et la soustraction, ou l'addition et le dessin des triangles. Il y a eu quatre paires d'activités semblables et quatre paires de dissemblables. On a testé deux fréquences de changement en faisant alterner les sujets une fois par minute ou dix secondes. On a constaté de grandes différences individuelles dans la capacité d'alterner les activités mais très peu de constance entre les tâches. On ne peut considérer que nos sujets alternent bien ou mal dans un sens général. Ils montrent très peu de constance dans leur capacité de faire des tâches variées comparées aux tâches constantes. Un sujet donné peut bien alterner dans une situation et mal alterner dans une autre. On a constaté que la capacité d'alterner dépend du type des tâches alternées et aussi il faut considérer si le changement a eu lieu entre des tâches semblables ou des dissemblables. Même en adoptant un critère peu précis de type, on a observé que les sujets ne tendent que très peu à bien alterner ou à mal alterner. En alternant des tâches semblables 38 sur 51 sujets ont mieux travaillé continûment; tandis qu'avec des tâches dissemblables 47 ont mieux alterné avec une activité alternante et seulement quatre avec une activité constant. La variation de la fréquence de l'alternation d'une minute à dix secondes a produit peu de différence dans les résultats pour l'activité alternante et pour l'activité constante.

Burri

INDIVIDUELLE UNTERSCHIEDE IN BEZUG AUF DIE FÄHIGKEIT, TÄTIGKEITEN ABZUWECHSELN

(Referat)

Es arbeiteten 50 Studenten beständig mit Addition, Subtraktion, Hinzufügung von Zeitwörtern oder Eigenschaftswörten an Hauptworter, Zeichnen von Dreiecken mit der Spitze nach oben oder nach unten, und Bildung von zwei Arten von Gestalten aus Gruppen bestehend aus je sechs Punkten. Dann wechselten die Vpp. zwischen zwei dieser Tätigkeiten ab—zum Beispiel, zwischen Addition und Subtraktion oder zwischen Addition und Dreieckzeichnen. Es gab vier Paare ähnlicher und vier Paar nicht-ähnlicher Tätigkeiten. Es wurden zwei Abwechslungsschnelligkeiten dadurch versucht, dass man die Vpp. in manchen Versuchen jede Minute und in anderen alle zehn Sekunden abwechseln lies. Es zeigte sich, dass Vpp. in ihrer Fähigkeit, Tätigkeiten abzuwechseln stark verschieden sind, obwohl sie von Versuch zu Versuch nur wenig Beständigkeit in Bezug auf diese Fähigkeit aufweisen. Die einzelnen Vpp. liessen sich nicht als im Algemeinen "gute" oder "schlechte Abwechsler" [good or poor alternators] beschreiben. Sie erwiesen nur wenig Beständigkeit in ihrer Fähigkeit, abwechselnde im Vergleich mit konstantbleibenden Arbeiten auszuführen. Eine gegebene Vp. kann sich in einer Situation als ein "guter Abwechsler" und in einer anderen Situation als ein "schlechter Abwechsler" darstellen. Man fand, dass die Fähigkeit, abzuwechseln, einerseits von der Art der abgewechselten Tätigkeiten abhängig war, und anderseits davon, ob die Abwechslung zwischen ähnlichen oder zwischen night-ähnlichen Tätigkeiten stattfand. Auch dann, wenn man einen sehr allgemeinen Massstab der Abwechslungsfähigkeit anwendete, erwies sich die Tendenz der einzelnen Vp., sichs als guten oder als schlechten Abwechsler zu zeigen, als sehr gering. Bei der Abwechslung zwischen ähnlichen Tätigkeiten arbeiteten 38 der 51 Vpp. besser ununterbrochen. Bei der Abwechslung zwischen nicht-ähnlichen Tätigkeiten arbeiteten 47 unter Abwechslung besser, und nur 4 besser ununterbrochen. Eine Änderung der Häufigkeit von einer Minute auf 10 Sekunden bewirkt nur wenig Unterschied in den Befunden in Bezug auf abwechselnde und beständige Tätigkeit.

BURRI

PSYCHOLOGY AND TECHNIQUES OF DISCOVERY*

T. N. Whitehead

From the Graduate School of Business Administration of Harvard University

Undoubtedly, experimental psychology has had a somewhat disappointing career. As a science it is still unformed and suffers from a difficulty in establishing a few major abstractions which could serve as a foundation for subsequent deductions. Newton's laws of motion are the classical examples of the sort of thing I mean and provide some foundation for the physical sciences.

Now until such time as a body of relevant abstractions—or major variables—is discovered in psychology, their search becomes one of the chief tasks of the investigator; this search dominates the literature of the subject, but the difficulty is to find some adequate techniques for the purpose. The exact sciences do not afford much assistance, for their best-known methods relate to the uncovering of relationships between known abstractions, rather than in searching for the latter themselves. This point is often overlooked, for it happens that the elaboration of scientific method did not make much progress until after the more exact sciences had accumulated a good working supply of useful abstractions. Bacon is usually quoted as the father of scientific method, but, beyond insisting on the necessity for experiment, he did little to elaborate the subject, and it did not receive adequate attention until about the middle of the nineteenth century.[1] From then on scientific method has been a favorite subject for discussion amongst scientists; but this discussion has mainly confined itself to those sciences in which both the type of abstraction and quite a number of actual abstractions were already known. True, the subject of new sciences was not wholly disregarded, but their essential peculiarity was not noticed and the methods recommended were all taken from the older sciences; thus the difference between two distinct stages of scientific growth has become blurred. In the first stage the chief need is to establish the type of abstraction

*Accepted for publication by W. J. Crozier of the Editorial Board and received in the Editorial Office, July 18, 1933.

[1]E.g., J. S. Mill published his *System of Logic* in 1843.

to be entertained, together with a few adequate examples; this at the same time helps to determine the subject itself and paves the way for the second stage. Here the relations between the known abstractions are investigated and new abstractions are uncovered in the process. The latter process will not be successful, however, unless the provisional abstractions are sufficiently adequate to hold the research more or less on fruitful lines. The danger of confusing these two stages in discovery is that a premature adoption of the methods suitable for the second may actually delay the progress of the first, by unduly narrowing the scope of investigation into channels which carry no guarantee of significant content, and thus prevent a broad survey of the field.

A rather interesting situation seems to have arisen in the science of physics, for there is evidence that a stage has been reached in which past abstractions are becoming inadequate, and modern speculations as to the required modifications do not suggest any very clear line of progress. Scientists are better equipped to investigate relations between abstractions than to uncover the basic abstractions themselves.

As an example of an experiment possible in a science with well-established abstractions, let us consider the procedure of a physicist who wishes to measure the pressure of the atmosphere adjusted to sea level. He begins by observing the height of a column of mercury whose weight is just balanced by the pressure of the atmosphere surrounding it. But this alone will not give him the required pressure; two other measurements are required. First, the temperature of his barometer; for both the mercury and the glass surrounding it vary in volume with temperature; then the height of the instrument above sea level, for height affects the density of the air. The first of these variables could be controlled, although it is far simpler to correct for it on paper; but, short of digging a pit right down to sea level and conducting observations at the bottom, the second variable must necessarily be handled by calculation.

The point to notice in this procedure is that the observation of the required fact, atmospheric pressure reduced to sea level, is dependent on a number of variables: height, temperature, and others which are not enumerated because they are held sufficiently constant, or because their variation does not affect the issue, although their mere presence is an essential part of the whole situation.

Moreover, the dependence of the given variable on the others is quite fundamental and cannot be conjured away by any process whatever, either in fact or on paper. What can be done is to admit all this and define a standard condition for the unwanted variables (standard height; sea level; standard temperature: 0° C.) and then to restate the actual observations (of atmospheric pressure) in terms of the "standard" conditions of the unwanted variables.

Thus the "abstraction" of given parts of the whole situation, as performed by scientific procedures, is in reality no isolation at all; on the contrary, it is a direct acknowledgment of the total situation coupled with a technique for discounting the latter's unwanted variations. A scientific abstraction is thus a simplification of reality in so far as all unwanted variables are supposed constant (not absent); or, alternately, their *variations* may be immaterial to the problem in question. Thus in considering the application of the laws of motion it is not necessary to consider the color of the bodies involved; neither, on the other hand, are the bodies supposed to be devoid of color.

Now, aside from possibilities of control and measurement, three things are needed to perform the operations outlined above:

1. Most obviously, laws connecting a number of variables must be quantitatively known.

2. As an antecedent condition to the above, the variables must be recognized as such in order that the needed laws shall have been sought and discovered. But nature does not present itself as a bundle of "variables" to naïve inspection; rather, it is a changing whole composed of endless detail. A variable is a high abstraction and implies a history of successful elimination and selection of relevant data. This is an end product of science, not its point of departure.

3. Finally, and most fundamental of all, it must be supposed that the final effect of the *changes* of a number of variables acting in concert is, at least ideally, deducible from a consideration of these changes taken separately. This is an act of faith, and demands a pragmatic justification.

It is clear that no science can have started with a knowledge of its own relevant abstractions and laws, and it is precisely the identification of useful variables, possessing simple and stable interrelationships, which constitutes the first and most difficult step in the

birth of a new subject; determination of exact relations between known variables dominates the second and easier stage.

The outstanding mark of inferior abstractions, elevated to the rank of major variables, is their hopeless inability to stand above the mass of surrounding detail and preserve stable relations to each other; moreover they are not even approximately indifferent to a useful proportion of each other's changes and no assumptions as to the constancy of the remainder will result in a stable relationship between any two. This was the position of chemistry until recently and is largely the position of the psycho-social studies to this day. The effect is to overemphasize the interdependence and complexity of the field.

A new science in its first stage of discovery seems to require rather broad methods for conducting preliminary surveys of a wide field. Obviously, there is no easy way of securing this, and research on this plane will continue to resemble a game of "hide and seek" until some more adequate procedures are evolved; though, judging from the progress of other sciences, this need not lead to undue pessimism. And, after all, the power of directing thought along the lines of a general search for important abstractions is by no means absent, although, for the most part, the process is very indefinite and should be classed as an "art" rather than a communicable technique. Perhaps in the nature of things this is to be expected; although a few procedures have advanced rather beyond this stage; and, as an example, one branch of mathematics seems to be peculiarly suitable for use in this connection. I refer to the mathematics of probability. For, to the extent of our ignorance of the useful abstractions—or major variables—capable of standing in stable relationships with each other, are we confronted with an overwhelming number of small variables; so numerous and so small that they cannot be disentangled in practice. Now the mathematics of probability is in essence a means for expressing those characters which are to be found in situations *so described* as to be influenced by a very large quantity of very small variables. I say "so described" advisedly, for no situation is ever subjected to mathematical analysis, but only an abstraction embodied in a description, and usually referred to as "the data."

The mathematics of probability has two distinct utilities in this connection. In the first place, it enables certain characters of these

small variables to be described en masse, as, for example, their probable combined effect in causing variations in the situation as described; this is the usual use of probability and need not be elaborated here. Its second use directly assists us to determine whether any given description does indicate the presence of a few major variables mixed up with the mass of minor ones. In detail the method varies somewhat with the circumstances, but the general lines are the same in all cases. First, a "statistical population" is described with respect to some one character; then a criterion is used to ascertain whether the variations of this character, amongst the members of the population, are of a type consonant with the theory that these variations are produced by a *very large number of very* small variables.[2]

If the answer be in the positive, no major variables are indicated. Should the answer be in the negative, however, this is a clear indication that the data can only be accounted for by the presence of one or more major variables; perhaps, though not necessarily, mixed up with a number of small variables.

The criterion is a very simple one: either a given distribution is compatible with the hypothesis that it is caused by nothing but a large number of very small variables—or it is not; in the latter case it follows that some few larger variables are present.

Perhaps one or two examples would help the argument at this stage; these will be chosen from work with which the present writer is connected. No attempt will be made to substantiate the following statements as a full account is to be published in due course.

For some years measurements have been automatically kept of the time taken by a few women, skilled operatives, to complete a repetitive assembly job which ordinarily takes slightly under a minute for its performance.[3] If a suitable number of consecutive assembly

[2]The more usual definition of variations due to "chance" is that they are produced by a "very large number of very small variables, *all acting in mutual independence.*" I have omitted the last condition in the text as it is really redundant. For if a number of small variables are not independent, they can always be described as one larger variable. The whole thing really hinges on a matter of definition.

[3]The research referred to was undertaken by the Western Electric Company at its Hawthorne Works, Chicago. The job consists of the assembly of small relays. The time taken to assemble every relay by each operative was measured with an accuracy of about 1/5 second; and the measurements extend continuously over several years.

times be taken from one operative's record, perhaps stretching over about a week, and plotted in the form of a frequency curve, it has the general appearance of a skewed probability curve with its mode towards the fast end. In this curve the abscissa represents assembly times, from about 40 to 90 seconds; the ordinate is marked off in percentages from 0 to about 7%, and these represent the percentages of the assembly times which fall within each second's interval on the base. In the case of one operative, the fastest assembly for a given period in 1928 was completed in about 40 seconds, a negligible number were assembled in over 90 seconds, whilst the mode falls at about 54 seconds with a frequency of about 6%.

The curve is clearly skewed, and this in itself indicates that some one or more large tendencies, or major variables, are present; but investigation shows, in addition, that no simple formula will express the form of this skew.

Luckily, for certain periods, the executive in charge of this experiment had recorded every assembly in which the operatives had experienced some quite unusual difficulty or unexpected happening.[4] It is thus possible to assemble some frequency curves from which one class of assembly times are eliminated; those which carry some gross interruption to the ordinary work procedure. These "accident-free" assembly curves are also skewed, but are well fitted by a class of curves known as "logarithmic probability curves." This strongly suggests that these curves, unlike the original ones, refer to homogeneous populations, although their departure from the form of normal probability indicates that some major variable still remains to be discovered.

Now, from a knowledge of the assembly job, it is reasonable to assume that variations in the assembly time of "accident-free" assemblies do not arise from any differences in the work involved, but rather from some variation in the way that this work is performed. With such an assumption we should expect the resulting frequency curve to be of the logarithmic variety. This cannot be discussed in detail here, but it depends on the way in which any increase of the average time taken to perform a given repetitive job will also increase the probable difference of times taken as between repetitions. Natur-

[4]E.g., when an operative had tried to assemble a defective part, or accidentally knocked over the partially completed assembly, or blew her nose, etc.

ally the various operatives do not all give the same curve, but they are all logarithmic in form.

A logarithmic probability curve differs from a normal curve in that one additional constant has to be determined before the form of the curve is definitely fixed. In the present application this constant is the time during which variations of assembly speed have been taking place.

It seems natural to suppose that, if an operative has completed a number of assemblies in an average time of, say, x seconds each, then, equally, the average time available for each assembly time to vary from the others is also x seconds. In other words, an assembly is subject to the chances of variability (as regards speed) over the whole time taken for its completion. This seems obvious, but it is not true. For an examination of the actual frequency curves, taken from the work of any of the operatives, brings out the unexpected fact that nearly the whole variation in assembly times takes place in about half the time it took to complete an average assembly. Thus, in the case of the operative already mentioned, assemblies were completed in times varying from 40 to 90 seconds with a maximum frequency of assemblies completed in about 54 seconds.[5] The analysis of this particular curve shows that in each relay some work was proceeding at a uniform rate for something like 25 seconds; the whole difference between 40 and 90 seconds being due to what was happening in the remaining 15 to 65 seconds. Of course, the constant work, occupying 25 seconds, may not occur in one lump; in fact, the probabilities are quite the other way. The most likely explanation, in the light of all available evidence, seems to be that the constant factor in the work relates to actual speed of movement, while the variable factor represents the skill with which actions are organized and performed. It happens that the actions required for assembly fall rather markedly into two categories. One consists of very easy gross movements, mainly of the arms; while the other is made up of small manipulative actions of the fingers and hands; these latter actions are markedly variable in the success which attends them and demand skill for their performance. In an average assembly these two types of work each occupy about half the total time.

[5]For reasons outside the present argument it is more convenient to quote the "mode" than any other central tendency.

This explanation cannot be taken as proved though there is evidence to support it, which will be given at the proper time; but, if the present analysis is faulty, some other will be required to account for the known departures of the actual frequency curves from that of normal probability forms.

This example shows the actual use of the criterion, discussed above, to uncover the presence of major variables acting on a statistical population. In the first instance the population was shown to be heterogeneous; in the second it was clearly acted on by some influence, or factor, which accounted for its skew;[6] finally, the failure of correspondence between the time of assembly and the time during which the assembly was liable to variation indicated the necessity for some further discrimination. In no case does the criterion indicate what the major variables are, although it both shows their presence and also provides some test as to whether any attempted interpretation is substantially correct. For, to test any hypothesis, it is only necessary to discount the effects of the hypothetical major variable and see whether the distribution, so rectified, follows the required form.

This proof of the presence of major variables has great value as a positive test, but less weight can be attributed to negative findings; for no logic can be applied to an actual situation, but only with some description, an abstraction from the totality; so the fact that

[6]Since writing the above I am indebted to Professor W. J. Crozier for reminding me that a regular skew in an otherwise normal distribution may well result from an unwise choice of variable. Thus, to use his example, if the distribution of variation in the diameter of otherwise similar pills should be found to vary normally, the corresponding distribution of volumes will, of course, be skewed, owing to the third-power relation connecting diameters and volumes of spheres. More generally, if some variable be distributed according to the normal law, only linear functions of that variable can also be so distributed. If a given variable is distributed according to a skew law, and it is found that the distribution of some function of this variable is normal, either the variable or its function may prove susceptible of carrying the more pregnant physical interpretation. In the first instance, the departure from normality falls to be interpreted; in the second, a different variable is accepted with its corresponding interpretation. In general, a variable and its non-linear function do not contain the same number of parameters, and call for rather different interpretations. Thus, in any case a departure from a normal distribution is clear warning that some further consideration of major variables is required, although the outcome may be a restatement of an old variable, rather than the addition of a new one.

no major variable is disclosed indicates only that the description provided does not contain positive evidence for its existence. Thus, a situation which, described in one way, appears as a typical "chance" phenomenon, described from another angle, may be found to contain important abstractions.

For instance, if the output of any two operatives, mentioned above, be separately examined, it will be found that the time taken to perform any one assembly is quite independent of that taken to complete the assemblies immediately before or after. In other words, when slow trends are eliminated, the actual order in which the various assembly speeds were arranged throughout a day is a "chance" order; no tendency is shown for small groups of fast or slow assemblies to fall together; nor, on the other hand, are fast assemblies immediately preceded, or succeeded, by slow assemblies in more than the probable number of occasions; assuming a chance arrangement of the known speeds for the day. This applies to any two operatives even if sitting next to each other. Suppose, however, that the output of these two operatives be so described as to emphasize the relative times *at* which they complete assemblies, then a completely different picture is presented. For, on certain days at least, it is found that the actual times *at* which they complete assemblies are not arranged in a "chance" order with respect to each other—the operatives are clearly pacing each other; but this does not upset the "chance" arrangement of each operative's output as first described above.

Innumerable examples could be given in which the mathematics of probability can be used to uncover the presence of major variables; and a comparison between this technique and those used to determine the relations between known variables shows the far greater sweep of the probability test as an instrument of search, just as it also shows its inferiority as a means for determining exact relationships.

There is, of course, no good reason for supposing that all phenomena could be so described as to be conditioned by nothing but major variables; but, since any failure of description will result in hiding existent major variables, and since in psycho-social phenomena "chance" situations are remarkably frequent, it seems reasonable to hope that they will be greatly reduced in the future by a better understanding of the type of simplification required in this field.

One word of caution, however, is needed in connection with the statistical technique outlined above. Although we have a criterion for the presence of major variables significant for the description under analysis, it is clearly true, not only that an unfortunate description will cover useful major variables, but also that there exists a real danger of elevating unwise abstractions to the rank of major variables. The type of major variable uncovered is largely determined by the characterization of the situation described, and the importance of the former depends on the wisdom of the latter. Mathematical technique can never substitute for intuition, and its findings can have no importance beyond that introduced by the fitness of the initial data. The mark of a worth-while major variable is to be sought in the simplification its use introduces in the description of relevant situations.

It will be seen that the use of the theory of probability indicated above has certain limitations and also great possibilities. The essential point is that, if the distribution of any given character in a population is not arranged comformably with the hypothesis that it is caused by nothing but very small variables, it follows that some few larger variables are present; if known to be present, these larger variables can usually be found and studied, for the chief difficulty in any search is to circumscribe the area for investigation, rather than to detect the thing sought when its immediate presence is known. Moreover, as already indicated, the same mathematical procedures provide a test for the adequacy of the variable sought.

A limitation to the uses of this technique lies in the necessity for collecting a population of sufficient size; this obviously precludes its use in many cases, for so much psychological data does not lend itself to repetition. Nevertheless, quite a volume of suitable statistical data has been, or could be, collected; for example, intelligence tests including trade tests, output and other data referring to industrial situations, and, in fact, any numerical data relating to tests or situations which are capable of being repeated with sufficient accuracy and frequency. Industrial situations have a peculiar importance as a field for the observation of human beings; partly because it is easier to obtain adequate statistics relating to industry than in other activities; but even more importantly because industry is favorable to such control of experimental conditions as is possible and desirable in the present state of psychology.

In the first place, industrial occupations represent an important part of human activity. In this respect they compare very favorably with laboratory experiments on individuals. In the latter case, there is very little guarantee that significant human situations are being studied, or that the individuals studied have not been thrown into highly exceptional mental attitudes by the conditions of the experiment itself. This would not be so important if the significance, and these abnormalities, could be assessed; but in the prevailing ignorance as to the important major variables all such assessment is little more than guesswork.

Then industry somewhat simplifies the human situation by introducing the usual controls relating to the behavior of employees without going so far as to create unusual situations; and this is greatly assisted by the prevalence of semi-repetitive work.

It is not suggested that the mathematics of probability represents the only technique available in the search for major variables; but it is undoubtedly an important instance, and a comparison between this and the techniques elaborated for the analysis of relations between known variables illustrates the essential difference between the two types of procedure. This suggests the advisability of sufficiently concentrating on the first type until a sufficient body of generalizations has been collected in the psychological field to give the second type its full scope.

Graduate School of Business Administration
Harvard University
Cambridge, Massachusetts

LA PSYCHOLOGIE ET LES TECHNIQUES DE LA DÉCOUVERTE

(Résumé)

La méthodologie de la science s'est élaborée principalement depuis l'établissement des sciences plus anciennes. Donc, les méthodes scientifiques les mieux connues s'appliquent plus spécialement à ces sciences où grand nombre de variables importantes sont déjà connues. Pour une science nouvelle il faut des techniques qui aideront à découvrir des variables utiles, ou abstractions, plutôt que des méthodes désignées à explorer leurs relations après qu'on les aura trouvées. On peut nommer cette première étape dans la fondation d'une science le stage de la découverte.

Pour la plupart, la découverte est un art plutôt qu'une technique communicable, bienque celles-ci ne manquent pas complètement. Un excellent exemple d'une technique, adaptée à aider la découverte des variables utiles

dans une science nouvelle, consiste en l'usage suivant des mathémathiques de la probabilité: On décrit une "population statistique" à l'égard d'un seul caractère; on teste les variations de ce caractère, parmi les membres de la population, pour voir si elles ont pu être produites seulement par un très grand nombre de très petites variables; sinon, une ou deux plus grandes variables, signifiantes pour la description donnée, sont probablement présentes. On discute l'usage de ce critère brièvement avec exemples.

WHITEHEAD

PSYCHOLOGIE UND TECHNIK DER ENTDECKUNG

(Referat)

Die Methologie der Wissenschaft ist hauptsächlich seit der Begründung der alten Wissenschaften ausgearbeitet worden. Daher beziehen sich die bestbekannten Methoden ganz besonders auf jene Wissenschaften, worin schon eine beträchtliche Zahl von wichtigen Variablen bekannt sind. Für eine neue Wissenschaft ist eine neue Technik notwendig, die in der Entdeckung von brauchbaren Variablen behilflich ist, oder sind eher Abstraktionen als Methoden notwendig, die zur Entdeckung ihrer Beziehungen bestimmt sind, nachdem sie gefunden worden sind. Dieser erste Schritt in der Grundlegung einer Wissenschaft mag als Entdeckungsstadium bezeichnet werden.

Meistens ist Entdeckung eher eine Kunst als eine mitteilbare Technik, obschon diese nicht ganz fehlt. Ein ausgezeichnetes Beispiel einer Technik, die in der Entdeckung brauchbarer Variablen in einer neuen Wissenschaft behilflich ist, besteht in dem folgenden Gebrauch der Wahrscheinlichkeitsrechnung: eine "Statistische Bevölkerung" wird mit Bezug auf irgend ein Merkmal beschrieben; man prüft die Variationen dieses Merkmales unter den einzeln Gliedern der Bevölkerung, um zu sehen, ob sie wahrscheinlich nur von einer sehr grossen Zahl von sehr kleinen Variablen hervorgerufen worden sind; wenn nicht, sind wahrscheinlich eine oder zwei grössere, für die gegebene Beschreibung bedeutungsvolle Variablen vorhanden. Man diskutiert den Gebrauch dieses Kriterions kurz mit Beispielen.

WHITEHEAD

THE INFLUENCE OF SYMBOLIC REHEARSAL UPON THE RETENTION OF A MAZE HABIT*[1]

From the Psychological Laboratory of Yale University

ROBERT S. SACKETT

STATEMENT OF THE PROBLEM

An attempt has been made in the present experiment to determine whether or not symbolic rehearsal, interpolated between learning and relearning, has a beneficial influence upon the retention of a maze habit. Two types or degrees of symbolic rehearsal will be considered. Thinking through or endeavoring to recall the pattern of the maze in ideational or verbal terms is one type. The other consists of attempts to draw the pattern of the maze.

It is well known that the rehearsal of memory materials is beneficial to their retention. There is no special problem here, since the material learned is of a symbolic nature to begin with. However, this is not the case with skills where overt movements, rather than processes of a more purely representative type, are dealt with.

It has been repeatedly demonstrated, by Warden (24), Husband (5, 6, 7), and others, that in learning a maze the human learner often acquires something in addition to the ability to execute a more or less coordinated pattern of movements. This "something more" is a knowledge or representation of the maze pattern in terms of verbal, visual, or kinaesthetic imagery or, perhaps, actual verbalization. Having gained such a representation, the question is whether rehearsal in the absence of the actual stimulus-response situation, by means of these symbolic processes, will aid the learner in performing the actual activity at a later date.

If symbolic rehearsal can be demonstrated to have a beneficial

*Recommended by Edward S. Robinson, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, April 10, 1933.

[1]This article is a part of a dissertation presented for the degree of Doctor of Philosophy in Yale University. The project was carried out under the direction of Professor Edward S. Robinson. The writer wishes to express his sincere appreciation to Dr. Robinson for his many suggestions and valuable assistance.

influence upon the retention of a motor habit, the importance of this finding for the general problem of the retention of skills is fairly obvious. At the same time this importance will, no doubt, be limited by the degree to which a given motor habit can be symbolically represented.[2]

Aside from an interest in the problem for its own sake, symbolic rehearsal is a factor to be considered in retention experiments with the maze (and perhaps other skills). A number of experimenters working with maze retention have considered this matter important enough to take steps to control it. For instance, Tsai (23, p. 4) said nothing to his subjects about relearning until a few days before the retention test. Husband (6, pp. 272-273) told his subjects that they were to come back "to learn *another* maze, which will probably not take more than a third of the time it took to learn this one." Webb (28, p. 13), Waters (25, p. 221; 26, p. 294), Waters and Ellis (27, p. 266), and McGeoch and Melton (12, pp. 404-405) all told their subjects to refrain from any attempt to rehearse or recall the pattern in the interval between learning and relearning. In addition, McGeoch and Melton questioned their subjects at the time of relearning and discarded those cases where there was any doubt concerning rehearsal. Kreuger (11) and Koch (10) apparently made no attempt to control this factor.

The importance of determining the influence of rehearsal is further shown by Tsai's suggestion (23, p. 27) that the difference found by him between the general shape of the retention curves for human and rat subjects may have been due to attempted recall during the interval on the part of the men.

The influence of symbolic rehearsal must be known in order to determine whether it should be considered seriously or whether it may be safely disregarded in retention experiments with the maze and with motor habits in general.

---

[2]It may be observed that the present problem is related to some of the experiments in transfer of training. For instance, by means of a lecture on refraction, Judd (8) attempted to impart information to children concerning skill in hitting a submerged target with a dart.

Renshaw and Postle (16) attempted, unsuccessfully, to shorten the learning of a pursuit task by giving detailed verbal instructions to a group of subjects concerning the best methods of going about the task.

While these studies are similar to the present one in the attempt to substitute symbolic processes for overt activity, they differ from it in several important respects. In the first place, the information is imparted prior to actual practice rather than through practice. Furthermore, the main interest was focused upon acquisition rather than on retention.

Apparatus and Procedure

Three groups of twenty subjects each learned a maze under standard conditions. Following this, each group was given a different set of instructions regarding rehearsal. Seven days later the same maze was relearned. The subjects were questioned as to their activities during the learning and relearning periods and as to those in the interval between the two periods.

The Maze. The maze used was of the high-relief finger type originally described by Miles (14). The path was constructed of No. 14 copper wire bent into staples of suitable length. The design was laid out on a three-ply board, 45.5 centimeters square. The staples were forced into holes of a size slightly smaller than that of the wire.

The particular pattern employed was a modification of that used by Spence and Townsend (22). The last few units were changed and added to so that all sections conformed to the T-principle. In addition, the cross-sections at the ends of the blind alleys were left off. The pattern covered an area of 25 by 23 centimeters. The units between choices as well as the blind alleys varied in length from 2.5 to 4.5 centimeters. (See Figure 1.)

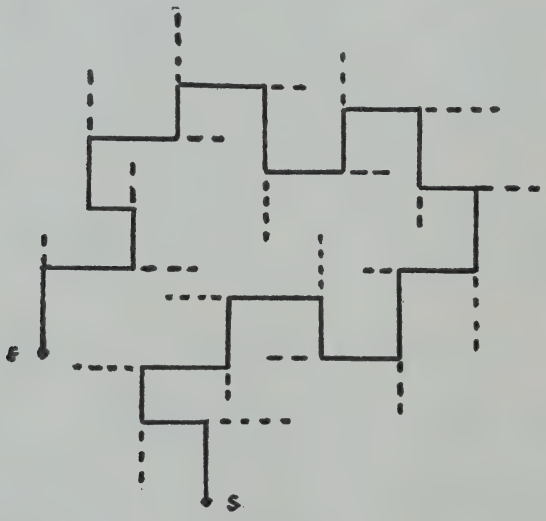


FIGURE 1
Diagram of the Maze Pattern
(One-fourth actual size)

The Miles standard blindfold technique consisting of taped goggles was used (14).[3]

The maze was kept in a constant position at the side of the table at which the subject was seated and was covered with a large black cloth when not in use. The learner was allowed to place his chair at the most comfortable distance from the table. However, care was taken that the median plane of the body was opposite the starting-point of the maze.

Reliability of the Maze

The reliability of the maze as computed by the method of total errors on odd *vs.* even trials was found to be +.93 with a standard error of ±.02. Considering total errors on the first half *vs.* those on the second half of the blind alleys, a reliability coefficient of +.67 with a standard error of ±.07 was obtained (Pearson coefficients). When these coefficients were corrected for halving by the Spearman-Brown formula and the standard errors calculated by Shen's formula the coefficients were, respectively, +.96 ±.009 and +.80±.05.[4] In every instance the number of cases was 60.

These values compare favorably with those obtained by other recent investigators with the maze and are large enough for the study of group differences.[5] However, it should be pointed out that all trials were included in the present computations. Consequently, as a result of the wider dispersion, the coefficients are higher than they would have been had only a certain section of the trials been used. This should be kept in mind when comparing these coefficients with those obtained from a certain set number of trials. Since all of the trials are used in arriving at our general results it seems wise to use them all in calculating the reliability.

---

[3]After placing a rectangular piece of crepe paper over the inside of the goggles, the writer found it advantageous to insert a snugly fitting wire ring into each eye-piece. This depressed the tissue against the lenses of the goggles and served to keep the paper away from the eyelids. Eighty per cent of the subjects kept their eyes open with this technique.

[4]Spearman-Brown formula: $r_x = \frac{Nr}{1 + (N-1)\,r}$ (4, p. 269)

Shen's formula: $\text{S.E.}_r = \frac{2\,(1-r)}{\sqrt{N}}$ (19)

[5]For some of the reliabilities obtained by other investigators see (1, 2, 3, 6, 7, 13, 15, 16, 18).

For a discussion of the methods of computing the reliability of the maze and the interpretation of the resulting coefficients see (21). (This article includes a fairly complete list of references to previous discussions.)

Subjects

The subjects employed were all men, university students averaging 20.2 years of age with a range from 17 to 25 years. The various ages were spread fairly evenly over the three groups. All of the subjects except four reported that they were without previous maze experience. The objective records as well as the verbal reports showed no reason for excluding these individuals.[6]

Learning Instructions

Before coming to the experiment, all of the learners knew that they would be needed on two separate occasions each of which would not last longer than an hour. They were not informed as to just when the second session would come, nor were they told that they were to learn a maze. Only after mastering the maze was it made known to them that they were to return a week later. They were informed that they would not be paid for either occasion unless they appeared on time for the second session. No records have been included where the time of relearning varied by as much as an hour from the time of learning the week before.

Upon reporting for the learning test, all subjects were warned not to say anything to anyone about the experiment. The goggles were then adjusted and the following instructions read to the learners:

> Your task is to learn to go from the entrance to the end of a raised-wire finger maze or labyrinth, without getting into any blind alleys or running off the wire path.
>
> There is but one correct path from the beginning of the maze to the end and this route is constant. In addition to the correct path there are a number of blind alleys where you will run off the wire.
>
> On each trial your object is to follow along the path, exploring different ones with your finger, in an attempt to reach the goal. The goal in this case is a little rubber button at the end of the maze. You will stop when you reach this point and put your tracing hand in your lap each time.
>
> An error will be counted for any entrance into a blind alley and for any attempt at retracing the correct path. You will

[6] It had been at least a year since these subjects had had any experience with a maze. None of them had ever learned a high-relief maze or one of multiple-T pattern.

be stopped when you start to retrace, but an error will be counted.

Your object is to learn to follow the correct path as quickly as possible without making any errors. The experiment will be completed only after you have made two runs in succession without an error.

Following the reading of these instructions, the experimenter placed the hand not used in tracing at one corner of the maze board. The index finger of the other hand was placed on a sample section on the right side of the maze board. The subject was told to run his finger all over this single T-section in order that he might know the nature of the path and the blind alleys. After he was told that this was a sample section, his finger was placed at the start of the maze. He was informed that he could keep his eyes either open or closed, just as he wished. He was further instructed that he would be paid for an hour whether it took him five minutes or an hour to learn the maze, but that the experiment would not be allowed to run over an hour.

Whenever a subject started to retrace, the experimenter told him that he was retracing and placed a pencil to prevent further movement back toward the start. With this technique, retracings usually disappeared after the first few trials and in any case were few in number. Retracings were recorded but are not included in this report.

Any entrance into a blind alley by more than a quarter of an inch was considered as an error. Time was taken with a stopwatch and was recorded in seconds for each trial.

The number and position of the errors were also recorded. Although the learners were told that they could rest between trials whenever they felt tired, only a few did so and then for not more than a minute or so.

Interval Instructions

After reaching the criterion of two errorless trials in succession, the individuals in the *drawing* group were asked to remove the goggles and to make five drawings of the maze with their eyes open. Each drawing was covered up before the next one was made. These instructions were then read to them:

You are to come back again next week on the same day

and at the same hour. In the meantime, draw the maze as often as possible just as you did now. Keep as accurate a record as possible of the number of times that you draw the pattern. Save all the drawings you can and bring them with you next week. After making a drawing, put it away, do not look at it when you make another one.

Each evening as you retire draw the maze five times just as you did today, never looking at the previous drawing when you make a new one. The whole success of the experiment depends upon your strict observance of these instructions. Do not wait until just before coming back to make your drawings.

After learning the maze, the subjects of the *thinking* group received somewhat different treatment. They were each given a sheet of paper on which a brief description of the visual, verbal, and motor methods of learning had been typed and were asked to classify themselves as to the method or methods used. If more than one had been employed, they were instructed to designate which method had been predominantly used. They were warned that the methods might easily be confused and to be as careful as possible in making their classifications.

The following instructions were then read:

You are to come back next week on the same day and at the same hour. At that time you will again try to go through the same maze without making any errors.

In the meantime, think of the pattern of the maze. Each evening before you retire, start at the beginning and think of the various moves you must make in order to get through without an error. Rehearse the pattern in this fashion five times each evening. This is the minimum; in addition, think your way through the maze as often as possible at other times.

Under no circumstances are you to draw the pattern or trace it in any way with your finger or by moving your head. The success of the experiment depends upon your strict observance of this warning.

Write down each day the number of times that you have thought your way through the maze and bring this account with you when you return next week.

The learners in the *non-rehearsal* group were treated in exactly the same way as those in the *thinking* group with the excep-

tion that the instructions given after they had made their classifications of learning methods were as follows:

> You are to come back next week on the same day and at the same hour. At that time you will again try to go through the same maze without making any errors.
>
> In the meantime, don't ever think of the pattern of the maze. Don't allow yourself to think of it. To do so will spoil the whole experiment. If the pattern should come to mind, think of something else at once.
>
> Under no circumstances are you to draw the pattern or trace it in any way with your finger or by moving your head. The success of the experiment depends upon your strict observance of this warning.

The instructions as to activities during the interval between learning and relearning were read twice to all three groups, and the subjects were then asked to repeat the main points. Preliminary experimentation indicated that such repetition was advisable if the directions were to be understood fully.

Relearning Instructions

Upon appearing for the relearning test the members of all groups were told the following:

> I should like to know just what you did since last week as far as the experiment is concerned. Whether or not you followed instructions does not matter as such. As a matter of fact, I expect that a number of the people won't follow the instructions strictly. The important thing now—and it is extremely important—is to know what you actually did do during the interval.

The members of the *drawing* group were then asked to make five more drawings of the maze pattern, covering up each one before beginning the next. After this the goggles were adjusted and these instructions read:

> Your task today is to relearn the same maze that you learned last week. You will remember that your object is to learn to follow the correct path as quickly as possible without making any errors. The experiment will be completed only after you have made three runs in succession without an error.

After the relearning, the subjects were given a brief oral description of the verbal, visual, and motor methods of attack. They were asked to classify themselves as to the method used in *learning,* giving the method predominantly used if more than one had been employed. The same questions were raised regarding the relearning; the subjects were cautioned to keep the two periods as distinct as possible. Several further questions were also asked concerning activities during the interval.

Upon their return to relearn the maze, the individuals in the *thinking* group were requested to tell the method or methods used in thinking of the maze pattern during the interval. They were then questioned on several other points.

The subjects in the *non-rehearsal* group were also interrogated as to what they had done during the interval.

From this point on the treatment of the *thinking* and *non-rehearsal* groups was exactly the same. The relearning instructions were the same as those given to the *drawing* group. After relearning, the members of the *thinking* and *non-rehearsal* groups reported the mode of attack used during the relearning. In this connection they were again referred to the brief written description of methods.

Differences in Experimental Conditions

The *thinking* group and the *non-rehearsal* group can be compared directly as they were subjected to the same conditions, with the exception of the instructions as to activities during the interval. The *drawing* group, however, was put through the experiment before the other two groups were planned. Several defects were discovered in the technique, so it was thought best to avoid these errors in the *thinking* and *non-rehearsal* groups at the cost of not having them strictly comparable to the *drawing* group.

The members of the drawing group were not told definitely that they would have the same maze when they returned, contrary to the treatment of the other two groups. However, the verbal reports showed that all but three of this group did expect to do the maze again. One of these three had not considered the matter, while the other two thought that they would merely be questioned about the drawings upon their return. Had this group been told definitely that they would be retested on the maze, it is possible

that their relearning performance would have been even better than it was.

Another difference is that the *drawing* group was not questioned at all until after relearning, while the other two groups were interrogated after learning as well as before and after relearning. The reports of the drawing subjects are thus open to the possible errors of forgetting and coloring of the report by intervening activities.

There is also the possibility that the questioning after learning and before relearning may have brought about some extra rehearsal for the *thinking* and *non-rehearsal* groups. On the other hand, this possible advantage may be offset by the fact that the individuals in the *drawing* group were favored by five rehearsals immediately after learning and directly before relearning.

Furthermore, the *drawing* and *thinking* groups differ in that the drawing subjects reported more rehearsals than did those in the *thinking* group.

Finally, the subjects in the drawing group admitted that they had thought of the pattern and traced it in the air with a finger at times other than when they were drawing it.

Thus the *drawing* group may be favored by more rehearsals, more immediate rehearsals, and thinking and tracing in addition to the drawing. The *thinking* and *non-rehearsal* groups may be aided by definite expectation of a retest and possible rehearsal during questioning after learning and before relearning.

Results—Relative Retention

Equation of Groups. The degree to which the groups were equated in learning ability is shown in Table 1, where the average learning scores for trials, errors, and total time are given for each group.[7]

TABLE 1
Average Learning Scores and Standard Deviations

Groups	N	Trials		Errors		Time in sec.	
		Av.	S.D.	Av.	S.D.	Av.	S.D.
Drawing	20	20.45	8.47	64.25	28.96	811.80	376.98
Thinking	20	21.70	7.84	61.20	26.61	803.60	253.75
Non-rehearsal	20	21.80	7.93	63.60	24.99	812.65	311.31

[7]All computations in all tables include the criterial trials.

While the equation is not exact, when all three measures are considered, the groups are probably nearly enough alike to be comparable. The same thing is true, with the exception of time for the thinking group, for the standard deviations of the groups. These are also to be found in Table 1.[8]

The average ages in years and the ranges of the ages for the groups are given in Table 2.

TABLE 2

AVERAGE AGE AND RANGE OF AGES IN YEARS

Groups	Av.	Range
Drawing	19.6	16.9-23.5
Thinking	20.3	17.9-26.9
Non-rehearsal	20.8	17.7-25.0

The *thinking* and *non-rehearsal* groups are closely comparable as to both averages and ranges. However, the *drawing* group averages a year younger and has a narrower dispersion because of restriction at the high end of the distribution. In view of the large individual differences in maze-learning scores, this lack of equality in age, particularly at this level, is probably not serious.

RELEARNING SCORES

In Table 3 we have the average relearning scores for the three groups when the criterion of relearning is two errorless trials in succession. The standard errors of the means are given also. The number of cases in each group is 20.

TABLE 3

AVERAGE RELEARNING SCORES AND STANDARD ERRORS: TWO ERRORLESS TRIALS

Groups	Trials	Errors	Time in sec.
Drawing	4.30± .31	3.90± .79	101.00± 7.89
Thinking	5.75± .48	6.70±1.13	151.65±16.06
Non-rehearsal	8.35±1.07	11.55±2.22	205.05±27.44

By all three measures the *drawing* group has better (i.e., lower) scores than does the *thinking* group. The same is true of the *thinking* group with respect to the *non-rehearsal* group.

[8]The coefficients of relative variability add nothing to the information given by the standard deviations.

Table 4 shows the differences between the groups with the standard errors of the differences.

TABLE 4

DIFFERENCES AND STANDARD ERRORS OF DIFFERENCES: RELEARNING—TWO ERRORLESS TRIALS

Differences	Trials	Errors	Time in sec.
T minus D	1.45± .57	2.80±1.38	50.65±17.89
N minus T	2.60±1.17	4.85±2.49	53.40±31.71
N minus D	4.05±1.12	7.65±2.36	104.05±28.53

For all measures the differences between the *thinking* and *non-rehearsal* groups are greater than those between the *drawing* and *thinking* groups. But, because of differences in the conditions affecting the *drawing* group (as was pointed out on pp. 384-385), we cannot say that a given amount of drawing rehearsal is relatively more effective than a similar amount of thinking rehearsal, although this may be true.

Table 5 presents the critical ratios for the comparisons of the three groups on all three measures.

TABLE 5

DIFF./S.E.$_{diff.}$ FOR RELEARNING—TWO ERRORLESS TRIALS

Differences	Trials	Errors	Time in sec.
T minus D	2.54	2.03	2.83
N minus T	2.21	1.95	1.68
N minus D	3.63	3.25	3.65

Only in the case of the comparisons of the *drawing* and *non-rehearsal* groups are the differences three times their standard errors. The chances that the true differences are greater than zero vary between 95 and 99.75 in 100 for the other comparisons. These probabilities are high enough to suggest that all of the differences are real ones.

It seems safe to conclude from these results that rehearsal in the form of drawing the pattern aids maze retention as measured by relearning after one week. There is a fairly high, although not statistically reliable, probability that rehearsal in the form of thinking about the pattern also has a beneficial influence upon retention. There is a similar statistical probability, qualified by the experi-

mental conditions, that drawing rehearsal is more beneficial to retention than is thinking rehearsal.

The relative retention was also computed with a relearning criterion of three errorless trials in succession, by calculating the average difference between learning and relearning, and in terms of percentage saved.[9]

With the more difficult criterion the results were essentially the same as with the criterion of two errorless trials. The average difference method with a relearning criterion of two errorless trials showed the same general relationships between the groups, but none of the differences was at all reliable. This was also true of the savings method. These other methods of computing retention give the same picture as the first, although somewhat less clearly.

Conclusions

Both *thinking through* and *drawing* a maze pattern are beneficial to retention, with drawing having the greater influence under the conditions of the experiment. This statement is supported by the fact that the relations between the mean retention values remain generally the same whether retention is measured by relearning scores with a criterion of two errorless trials, relearning with a criterion of three errorless trials, the difference between learning and relearning, or percentage saved.

Although the difference between the groups is significant statistically only for the differences between the *drawing* and the *non-rehearsal* groups, by two ways of measuring retention, there is a reasonable presumption that they are all real differences.

Unfortunately, the question as to how much thinking equals a given amount of drawing or vice versa cannot be answered, because of the differences in the treatment of the two groups.

Results—Analysis of Drawings and Verbal Reports

Drawing Group. It will be remembered that the members of the *drawing* group were required to draw the pattern of the maze five times immediately after learning and just before relearning. Each drawing was covered up before another was made. We thus have

[9] By the formula: $\text{Savings} = 100 - \frac{\text{Relearning}}{\text{Learning}} \times 100.$

an objective record of the subjects' conceptions of the pattern before and after the interval of rehearsal. An analysis and comparison of these initial and terminal drawings may give some insight as to what happened during the interval.

Correctness of reproductions. Of the five patterns drawn immediately after learning those of only three individuals were all correct. Eight had none of them accurate.[10] However, only two subjects had all of the drawings before relearning correctly drawn; thirteen had no accurate reproductions.

The average number of correct drawings after learning was 1.6, while before relearning it was reduced to .55. This difference of 1.05 correct reproductions is 1.98 times its standard error. There are 98 chances in 100 that the true difference is greater than zero, which indicates a fairly strong but not statistically significant probability.

Two things appear from these results: In the first place, only 15% of the subjects were able to reproduce the pattern correctly five times immediately after learning. This suggests that most learners do not have an accurate symbolic representation of the maze even after going through it twice in succession without an error. Secondly, there was a tendency for the subjects to have a less accurate representation of the pattern after drawing it for a week than they had immediately after learning. In spite of the erroneousness of this rehearsal, it enabled them to relearn the maze more effectively than was the case with non-rehearsing subjects.

Variability of reproductions. The average number of different patterns or versions of the maze found in the five drawings immediately subsequent to learning was 2.55. The drawings before relearning show that the group averaged only 1.25 versions. There was a reduction of 1.30 versions after a week's interval of drawing. That this reduction is decidedly significant is shown by the fact that the difference is 4.8 times the standard error of the difference. That this effect is widespread is demonstrated by the finding that only four individuals drew all five patterns the same after learning while sixteen had all five alike before relearning.

We may conclude that there is a strong tendency for subjects

[10]A drawing was considered as correct if all of the twenty-five sections were present and in their proper sequence. The reproductions were not scored on the proportions of the sections.

to be less variable in their reproductions of a maze after a week of drawing the pattern than immediately after learning. It seems probable that the drawing rehearsal is accompanied by a settling upon one of the several possible symbolic representations as the most accurate one. The verbal reports of the subjects are in general agreement with the findings from the objective records.

Changes in length of reproductions. When correctly drawn, the maze pattern is made up of 25 units of path. The average number of units found in the first one of the five drawings after learning was 25.45. The number in the last of these five reproductions was 24.20, a reduction of 1.25 sections. This difference is 1.74 times its standard error, showing that there are 96 chances in 100 that the true difference is greater than zero. It may be concluded that there is a noticeable but statistically unreliable tendency for simplification of the representation of the maze in the last of the first five drawings after learning as compared to the first.

When the average number of units in the five drawings after learning is compared with that of the five before relearning, practically no difference is found. The averages are respectively 24.77 and 25.13, with a difference of .36 units. The difference is not as large as the standard error.

Correlation between number of rehearsals and number of relearning errors. The average number of drawings per person for the *drawing* group, including the five subsequent to learning and those prior to relearning, was 67.5. The median was 55 and the range 45 to 130.

It was thought that there might possibly be some correlation between the reported frequency of rehearsal and the number of errors made on relearning. However, the rank correlation was .03, indicating that there is no relationship; or, if there is, it is obscured by some other factor or factors, such as individual differences in relearning ability.[11]

This finding may make relatively unimportant the fact that the drawing group had more drawing, thinking, and tracing rehearsals than the thinking group. The same thing is true of the possibility of rehearsal on the part of the thinking group during the questioning

[11] By the formula: $R = 1 - \frac{6 \text{ sigma } G}{N^2 - 1}$ (4, p. 193).

after learning and before relearning. This may leave the drawing group favored by more immediate rehearsal and handicapped by the lack of definite expectation of a retest. On the whole, the thinking and drawing groups are probably more comparable than they would at first appear to be.

Furthermore, no significant correlation was found between number of errors on learning and the reported amount of rehearsal. The rank coefficient was .04. It was possible that a subject who did poorly on learning might try to make up for it by more frequent rehearsal, but there is no evidence for this. Since the individuals did not definitely know either their own scores or those of the other subjects, there was no basis for comparison to enable them to know how effective their learning was.[12]

It should be mentioned here that, in addition to drawing, nine subjects reported that they thought of the pattern at other times. One individual thought of the pattern fifty times aside from drawing it, while the others thought of it from two to ten times. Ten subjects said that they traced the pattern in the air with their fingers. The frequency of this varied from one to fifteen times. While these facts must be kept in mind when the drawing group is compared with the others, it should also be remembered that the reported frequency of rehearsal was found to have little influence upon relearning scores.

Thinking Group

Correlation between frequency of rehearsal and number of relearning errors. The average number of reported rehearsals for the *thinking* group was 38.3, with a range from 18 to 53. The median was at 39.5. The *thinking* group averaged 29.2 fewer rehearsals than the *drawing* group. Even when the two sets of five rehearsals after learning and before relearning are subtracted from those of the drawing group, the difference is still 19.2. The fewer rehearsals on the part of the *thinking* group may be due in part to some slight difference in the wording of the instructions. Also, this

[12]An effort was made to correlate the number of errors made on the final five drawings with that made in relearning. However, no fair objective method of scoring the drawings could be devised. Finally, a rough method of scoring was used, which nevertheless differentiated the individuals fairly well. The Pearson coefficient for errors in drawing and in relearning was only +.02, again showing a highly insignificant degree of relationship.

difference may in part result from the greater concentration needed to rehearse the maze in purely symbolic terms, although there is no experimental evidence on this point.

When reported number of rehearsals was correlated with number of errors on learning the rank coefficient was .08. The correlation between frequency of rehearsal and number of errors on relearning was .02. Contrary to what might be expected, there appears to be no significant relationship between number of rehearsals and errors on learning and relearning in either the drawing or thinking groups.[13]

Verbal reports. None of the subjects in the thinking group admitted drawing or tracing the pattern, although nine reported that they were tempted to draw the pattern, and fourteen said that they felt some urge to trace the pattern out with a finger.

Non-Rehearsal Group

Verbal reports. None of the members of the non-rehearsal group admitted thinking his way through the maze, although seventeen reported that the pattern came to mind but that they immediately turned their attention to something else. Three subjects said that they were tempted to trace out the pattern in the air but did not do so. Only one reported any temptation to draw the maze. There is, of course, no way of telling how well the subjects actually obeyed instructions. All that can be said in this connection is that pains were taken to avoid any incentive for falsification.

Discussion

Our results give fairly reliable evidence that symbolic rehearsal does have a beneficial influence upon retention. While such a conclusion might have been expected, several considerations will show that this result could not have been taken for granted prior to experimentation. For instance, it might have been impossible to have gained an adequate symbolic representation which could have been used for purposes of rehearsal. Furthermore, granting

[13]It is possible that the frequency of rehearsal may be negatively correlated with the error score on relearning up to a certain number of rehearsals. Beyond this point, which probably varies for the individual, the frequency may have a negligible effect upon retention. The distribution of rehearsal and especially the immediateness with which it follows learning may also be important factors. The writer is at present gathering data upon the influence of the number of rehearsals upon retention under better-controlled conditions.

such a representation, it might have been or might have become so erroneous that its rehearsal would have brought about less effective relearning than no rehearsal at all.

The discovery that the number of rehearsals is not an important factor in determining relearning scores raises several questions which are open to experimental solution. One problem is that of the number of rehearsals which is optimal from the standpoint of relearning. As was mentioned before, there may be a certain number of rehearsals beyond which it is useless to go. It may be that rehearsal beyond a certain point is actually detrimental to relearning. A further problem is that of the most favorable distribution of rehearsal in the interval between learning and relearning.

These considerations bring up another point with regard to the limits of symbolic rehearsal. The writer is inclined to believe that the influence of symbolic rehearsal is limited to those skills in which there is a symbolic control of the movements involved. It seems doubtful that there is this control in the case of such skills as ball-tossing and mirror-drawing. Snoddy (20, pp. 40-47) found that in the period between mirror-tracings there is often visual and manual-motor imagery, but this was shown to have no value in controlling the actual movements. With a skill of this kind it is unlikely that symbolic rehearsal would be effective.

Although concerned with transfer of training and acquisition rather than retention, the experiments of Judd (9) and Renshaw and Postle (17) are relevant here. Judd found that a group of boys who were given a full theoretical explanation of refraction made no better scores than an uninstructed group in hitting a submerged target with a dart. But when the target was subsequently placed at a different depth, the instructed group excelled. No details are given concerning the nature of the two groups, and it is not stated whether or not the groups were equated in any way. There is, therefore, some question as to whether the difference in ability is due to the explanation of refraction or to some other factor. However, a part of the failure of the informed group to do better than the other group might be explained as due to the fact that there was no symbolic control over the movements involved. Not until such control was established by practice did the influence of the instructions make itself apparent. A part of the failure of Renshaw and Postle's fully instructed subjects to do well on the pursuit

task might also be explained as resulting from the lack of symbolic control of the necessary movements.

Thus the conditions under which symbolic representation of a motor habit is possible, as well as those under which rehearsal in these terms is effective, are problems for further experimentation. This applies not only to retention, but also to cases where the rehearsal occurs in the interval between practice periods in learning.

It should also be mentioned that 80% of the subjects in the thinking group rehearsed the pattern in predominantly visual terms. The remaining 20% employed the verbal mode. To the writer it seems important that rehearsal can be carried on effectively not only in verbal, but also in visual, terms.

Finally, our results show that under certain conditions symbolic rehearsal does appear to have a beneficial influence upon the retention of a motor habit. Whether or not rehearsal, such as may occur in retention studies where no specific instructions are given concerning rehearsal, is effective is a matter for further experimentation. If such rehearsal should prove to be effective the determination of the best means of preventing it is also a problem for research.

Summary and Conclusions

Three groups of twenty subjects each, approximately equated in age and learning ability, learned a high-relief finger maze under the same conditions. Subsequently, each group was given a different set of instructions regarding rehearsal. After an interval of seven days the groups relearned the maze under comparable circumstances. One group was told to draw the maze pattern as often as possible during the interval and was required to make five drawings immediately after learning and five directly before relearning. A second group was instructed to think through the pattern as often as possible during the week but not to draw or trace the maze. The third group was requested not to draw, trace, or think through the pattern. Verbal reports concerning the activities during the learning, relearning, and rehearsal periods were taken. These reports included the modes of attack used.

The following conclusions were drawn from the results, with the caution that in some cases they are based upon a very small number of subjects.

1. Symbolic rehearsal in the form of drawing is beneficial to

the retention of a maze habit after one week when retention is measured by several methods.

2. When the rehearsal is in the form of thinking through the pattern, it also appears to be beneficial. Although the difference between the experimental and control groups is not statistically reliable, there is some supporting evidence for this statement.

3. Drawing rehearsal seems to be more effective in its influence upon retention than is thinking. However, the difference is not reliable and may be influenced by the lack of equation of the groups in the immediateness of rehearsal and the definite expectation of a retest.

4. Most of the drawing subjects appear to have an inadequate representation of the maze pattern even immediately after reaching the criterion of learning.

5. There is a noticeable but statistically unreliable decrease in the accuracy of reproductions of the pattern after a week of drawing rehearsal.

6. A highly reliable decrease in the number of different patterns was found in the reproductions after an interval of rehearsal in the form of drawing.

7. An unreliable but apparent simplification of the pattern was found in the drawings just after learning.

8. No significant difference was found between the length of the graphic representations of the pattern before and after the interval of rehearsal.

9. Contrary to what might have been expected, neither type of rehearsal showed any correlation between reported frequency of rehearsal and number of errors in either learning or relearning. It is possible that attenuating factors may obscure some relationship. This result makes the differences in retention between the thinking and drawing groups more reliable than they at first appear to be.

10. Eighty per cent of the subjects of the thinking group rehearsed the pattern in visual terms. The remaining 20% employed the verbal mode.

References

1. Allison, L. W. Difficulty as a factor in the standardization of a maze. *J. Gen. Psychol.*, 1931, **5**, 514-518.

2. ————. The effect of variation of length of a high relief finger maze upon rate of learning. *J. Genet. Psychol.*, 1931, **34**, 403-407.

3. Barker, R. G. A temporal finger maze. *Amer. J. Psychol.*, 1931, **43**, 634-637.
4. Garrett, H. E. Statistics in psychology and education. New York: Longmans, Green, 1926. Pp. 320.
5. Husband, R. W. Human learning on a four-section elevated finger maze. *J. Gen. Psychol.*, 1928, **1**, 15-28.
6. ————. A note on maze learning with the time factor constant. *J. Gen. Psychol.*, 1929, **2**, 366-369.
7. ————. Analysis of methods in human maze learning. *J. Genet. Psychol.*, 1931, **34**, 258-278.
8. ————. Comparative behavior on different types of mazes. *J. Gen. Psychol.*, 1931, **5**, 234-244.
9. Judd, C. H. The relation of special training to general intelligence. *Educ. Rev.*, 1908, **36**, 28-42.
10. Koch, H. L. The influence of mechanical guidance on maze learning *Psychol. Monog.*, 1923, **32**, No. 147. Pp. 112.
11. Kreuger, W. F. C. Further studies in overlearning. *J. Exper. Psychol.*, 1930, **13**, 152-163.
12. McGeoch, J. A., & Melton, A. W. The comparative retention values of maze habits and nonsense syllables. *J. Exper. Psychol.*, 1929, **12**, 392-414.
13. McGinnis, E. The acquisition and interference of motor habits in young children. *Genet. Psychol. Monog.*, 1929, **6**, 203-311.
14. Miles, W. The high relief finger maze for human learning. *J. Gen. Psychol.*, 1928, **1**, 3-14.
15. Nyswander, D. B. A comparison of the high relief finger maze and the stylus maze. *J. Gen. Psychol.*, 1929, **2**, 273-289.
16. Peterson, J., & Allison, L. W. Effects of visual exposure on the rate and reliability of stylus maze learning. *J. Gen. Psychol.*, 1930, **4**, 36-48.
17. Renshaw, S., & Postle, D. K. Pursuit learning under three types of instructions. *J. Gen. Psychol.*, 1928, **1**, 360-367.
18. Scott, T. C., & Nelson, B. B. Factors affecting the reliability of the maze: a comparison of the high relief finger maze and an improved form of the stylus maze. *J. Gen. Psychol.*, 1932, **6**, 70-89.
19. Shen, E. The standard error of certain estimated coefficients of correlation. *J. Educ. Psychol.*, 1924, **15**, 462-465.
20. Snoddy, G. S. An experimental investigation of trial and error learning in human subjects. *Psychol. Monog.*, 1920, **28**, No. 124. Pp. 81.
21. Spence, K. W. The reliability of the maze and methods of its determination. *Comp. Psychol. Monog.*, 1932, **8**, No. 45. Pp. 45.
22. Spence, K. W., & Townsend, S. A comparative study of groups of high and low intelligence in learning a maze. *J. Gen. Psychol.*, 1930, **3**, 113-130.
23. Tsai, C. A comparative study of retention curves for motor habits. *Comp. Psychol. Monog.*, 1924, **2**, No. 11. Pp. 29.
24. Warden, C. J. The relative economy of various modes of attack in the mastery of a stylus maze. *J. Exper. Psychol.*, 1924, **7**, 243-275.

25. Waters, R. H. The influence of large amounts of manual guidance upon human maze learning. *J. Gen. Psychol.*, 1930, **4**, 213-228.

26. ————. The effect of incorrect guidance on human maze learning. *J. Comp. Psychol.*, 1931, **12**, 293-301.

27. Waters, R. H., & Ellis, A. L. The relative efficacy of free and guided learning when equated in terms of time. *J. Comp. Psychol.*, 1931, **12**, 263-277.

28. Webb, L. W. Transfer of training and retroaction. *Psychol. Monog.*, 1917, **24**, No. 104. Pp. 90.

New Jersey College for Women
New Brunswick, New Jersey

L'INFLUENCE DE LA RÉPÉTITION SYMBOLIQUE SUR LA RÉTENTION D'UNE HABITUDE DE LABYRINTHE

(Résumé)

On a fait cette expérience dans le but de déterminer si la répétition symbolique d'une habitude motrice, interpolée dans la période entre l'apprentissage et le réapprentissage, aide la rétention après sept jours.

Trois groupes, chacun de vingt sujets, rendus égaux à l'égard de l'âge et de la capacité à apprendre, ont appris et réappris un labyrinthe dans des conditions comparables. On a dit à un groupe de dessiner la forme du labyrinthe le plus souvent possible dans l'intervalle entre l'apprentissage et le réapprentissage. Cinq des dessins ont été faits immédiatement après l'apprentissage et cinq immédiatement avant le réapprentissage. On a dit à un autre groupe de penser à la forme du labyrinthe le plus souvent possible dans l'intervalle, mais ni de dessiner ni de tracer sa forme. On a dit au troisème groupe de ne pas penser à la forme dans l'intervalle. On a obtenu des rapports verbaux des activités pendant les périodes de l'apprentissage, du réapprentissage, et de la répétition.

On a constaté que la répétition du dessin aide la rétention d'une manière définie. Ni l'un ni l'autre type de la répétition n'ont montré aucune corrélation entre la fréquence rapportée de la répétition et le nombre d'erreurs dans le réapprentissage. On discute brièvement les conditions dans lesquelles la répétition symbolique des habitudes motrices peut aider. On a obtenu des rapports verbaux des activités pendant les périodes de après l'apprentissage et immédiatement avant le réapprentissage.

Sackett

DIE EINWIRKUNG DER SYMBOLISCHEN WIEDERHOLUNG AUF DIE BEIBEHALTUNG EINER GEWOHNHEIT BEI DER BEMEISTERUNG EINES LABYRINTHES

(Referat)

Es war die Absicht dieser Untersuchung, zu bestimmen, ob die symbolische (ideologische) Wiederholung einer motorischen Gewohnheit, in die Periode zwischen dem ursprunglichen Lernen und dem Wiederlernen eigeschaltet, die Beibehaltung nach sieben Tagen fördert.

Drei Gruppen, bestehend aus je zwanzig Versuchspersonen, die in Bezug auf Alter und Lernfähigkeit gleichgeschaltet worden waren, lernten und wiederholten die Durchforschung eines Labyrinthes unter vergleichbaren Umständen. Eine Gruppe wurde ersucht, so oft wie möglich die Gestaltung des Labyrinthes zu zeichnen, während des Zeitintervalles zwischen dem Lernen und dem Wiederlernen. Fünf von den Zeichnungen wurden gerade vor dem Lernen und fünf gerade von dem Wiederlernen gemacht. Von den Mitgliedern einer zweiten Gruppe verlangte man, dass sie sich das Labyrinth während des Intervalles so oft wie möglich durchdenken, aber dass sie die Gestaltung [pattern] des Labyrinthes nicht zeichnen oder abzeichnen [trace] sollten. Die dritte Gruppe wurde ersucht, während des Intervalles nicht an die Gestaltung des Labyrinthes zu denken. Es wurden wörtliche Protokolle [verbal reports] gesammelt, über die Tätigkeiten der Versuchspersonen während der Perioden des Lernens, des Wiederlernens, und des Wiederholens.

Es zeigte sich, das eine Wiederholung durch Zeichnen des Labyrinthes eine bestimmt günstige Einwirkung auf die Beibehaltung ausübte. Der günstige Einfluss des Durchdenkens war weniger ausgeprägt. Weder die eine noch die andere dieser beiden Arten der Wiederholung lieferte eine Korrelation zwischen der gemeldeten Häufigket der Wiederholungen einerseits und der Zahl der Fehler bei dem Wiederlernen anderseits. Es werden die Bedingungen, unter denen eine symbolische Wiederholung motorischer Gewohnheiten günstig einwirken kann, kurz besprochen. Es werden auch die Befunde einer Analyse der Zeichnungen sofort nach dem Lernen und eben vor dem Wierderlernen gegeben.

SACKETT

AN ATTEMPT TO MEASURE CERTAIN PHASES OF SPEECH*[1]

From the Department of Psychology of Washington University

JOSEPH W. HAWTHORNE

HISTORICAL INTRODUCTION

The first systematic attempt to devise specific tests of speech was made by Stinchfield (4, 5). Previous measures of speech are reviewed and criticized by her in the references cited.

The Stinchfield tests consist of two parts. Part one is a subjective rating of behavior and speech reactions. Part two consists of objective measures of speech ability and includes (1) Articulation Test A, measuring ability to sound consonants; (2) Articulation Test B, measuring ability to sound vowels; (3) Spontaneous Speech Test, or the number of words used per minute in responding to pictures; (4) Oral Reading Test, or words read orally per minute; (5) Silent Reading Test, or the rate of silent reading; (6) Percentage of Relevant Words in Spontaneous Speech, or the number of circumlocutions used; and (7) a vocabulary test.

Stoddard (6) used a series of articulation tests involving reading, recognition of sounds made by the experimenters, and the naming of pictures, as tests of speech in the Detroit City Schools.

With such excellent tests already in the field, the following study might be considered superfluous were it not for the fact that it seemed possible to draw several conclusions from this study that either were not or could not be drawn from the previous studies in the field.

*Recommended for publication by John Paul Nafe of the Editorial Board and received in the Editorial Office, April 28, 1933.

[1]Much credit is due Miss Maude Sewell for the collection of much of the data for this study, to Miss Melba Buchmann for a great deal of clerical help, and to a number of other students of the author who willingly helped to check over the computations involved in the work.

Aims of This Study

It was the primary purpose of this study to devise a means of objectively rating the degree of speech difficulty of any particular individual. It was hoped that such a measure would provide a useful tool in individual diagnosis, in determining the effectiveness of various therapeutic procedures, in the further study of speech impairment and its relationships with other factors, and in such other instances in which a measuring device is found useful. A secondary aim was the incorporation in the tests of such material as would furnish an individual diagnosis of specific defects in speech. Thus, it was hoped that the tests would quantitatively measure the amount of speech impairment in comparison with established norms and at the same time would qualitatively determine the nature of that impairment.

Procedure

A battery of tests was accordingly devised, which, it was hoped, would involve all of the aspects of speech that could be put into a test of this nature. For this purpose the following abilities were tested:

1. The ability to read the printed page orally
2. The ability to name familiar objects from pictures
3. The ability to speak spontaneously in answer to certain definite and simple questions

It is freely admitted that, if a comprehensive survey of one's ability to speak were to be made, the above would have to be supplemented by some measure of the ability to speak in the home, in company of friends, strangers, before public gatherings, and the like. The inclusion of measures of such abilities in a test of this nature seemed inexpedient, however, and it is believed that those abilities which were finally included are sufficient for at least a fairly adequate measure.

It is also admitted that, at least in the tests involving the reading of words and the naming of objects, there is grave difficulty that some ability other than speech is being measured, such as reading ability or general intelligence. That such difficulty has been at least partially obviated, it is hoped will appear in the body of this study.

Construction of the Tests

In the construction of the reading speech tests the following limitations were borne in mind:

1. The vocabulary of the material should be so simple that a minimum of reading ability or intelligence would be measured.
2. The material should be selected so as to be of maximum interest to the child.
3. The material should contain a maximum of English sounds.
4. Printing, where used, should be of optimum readability.

In devising this test the Bobbs-Merrill Primer (1), and the Bobbs-Merrill First Reader (2) were used as a vocabulary source. Many different types of stories were tried but those dealing with speaking animals were, for some reason or other, found to be the most interesting to the younger children of the group used for this study. The two finally selected were printed on separate cards and were as follows:

TEST I, PART I

Once there was a little boy. His name was little black Sambo. He had a red coat, blue trousers, and red shoes. He went for a very long walk in the forest. A fox jumped out and said, "Oh, I am going to eat you."

"Please do not eat me, Mr. Fox. Why do you not eat the goat or her four children?"

TEST I, PART II

A fox fell into a well. He did not know how to get out. The water was very cold and deep. Soon a goat came by. She looked into the well and saw the fox.

"Oh, come down and taste the water. It is fine and cold."

So the goat jumped into the well. The fox, jumping upon her back, soon got out of the well.

"Good-day, goat. Take all the water you want."

Then the fox ran away to his little fox children.

Test I, Part I contains all of the English articulations except the blends. Test I, Part II contains all of the English articulations except the blends and the vowel difficulties "oi," "ur," and "wh."

With much the same limiting factors in mind, the second test was constructed by assembling from the Sears Roebuck catalogue

a series of uniformly sized pictures of objects familiar to children, the names of which contained all the vowel and consonant sounds used in English, in initial, medial, and final positions. These pictures were mounted on 8½" x 11¾" cards in three rows of three pictures each, uniformly spaced. The cards were bound together in the form of a booklet which consisted of eleven pages of cards including the sample card on the first page. Following is an outline of the picture booklet together with the sounds which the pictures are intended to elicit.

In order to devise an adequate test of spontaneous speech, a list of 25 questions was prepared, all of which were orally put to a group of children in the lower grades of the elementary school. From this list were chosen those three questions to which the responses seemed the most natural and free, and which contained the greatest number of words.

Card I Sample		
chair	ironing board	ice skates
iron	tool chest	cowboy suit
pencil	rattle	glove

Card II		
H hammer	H rocking horse	Wh wheel barrow
F foot ball	F coffee pot	F golf club
V violin	V gloves	V glove

Card III		
Th thread	Th nothing	Th cloth
S stove	S bicycle	S horse

Card IV		
Z zebra	Z scissors	Z dishes
Sh ship	Sh machine	Sh flash light
R rocking chair	R horns	R bear

FIGURE 1
TEST II—PICTURE ARTICULATION TEST

Card V		
Ch chair	Ch kitchen cabinet	Ch watch
J jumping jack	J engine	J cage
L lamp	L soldiers	L ball

Card VI		
P piano	P type writer	P pipe
B black board	B table	B crib
M marbles	M umbrella	M drum

Card VII		
T train	T auto	T bat
D doll	D teddy	D sled
N nut	N ornament	N pen

Card VIII		
K kitty	K rocking horse	K desk
G gun	G buggy	G dog
W wagon	W towel	W ring

Card IX		
Ā table	A wagon	Aw auto
Oo skooter	O sewing machine	O top
T knife	I Indian	Oi cowboy

Card X		
Ee tree	E beads	E telephone
	U umbrella	yellow
Ou cowboy	bird	fireman

Card XI		
Br broom	Cr crib	Dr dresser
Fr freight	Tr truck	Gr green
Sp spoons	St stove	Sk skate

Card XII		
Sw swing	Sl slide	St street car
Pl plane	Bl block	Fl flag
Gl glove	Cl clock	X(ks) box

FIGURE 1 (*continued*)

The questions which were chosen to make up Test III were as follows:

"What did you have for breakfast this morning?"
"What did you have for dinner last night?"
"What game do you like best and how do you play it?"

Administration and Scoring of the Tests

Test I. Test I, Part I is administered with the following directions to the child:

> Here is a story. Read it aloud as quickly and as clearly as you can. Do you understand? Go.

In the meantime a card is handed to the child on which is printed in large-sized, well-spaced type the story comprising this part of the test.

Test I, Part II is administered in the same manner as soon as Part I is completed.

In scoring this test the examiner has in front of him a sheet of paper on which both stories are printed and which contains blanks for recording the name of the subject, his age, grade, and whatever other information is considered pertinent. Total time of reading and total errors are recorded for each test and, in addition, the following record is obtained while the subject is reading:

1. Words mispronounced or unpronounceable are underscored.
2. Words inserted are indicated by a circumflex.
3. Words or parts of words omitted are encircled.
4. Words in which stuttering occurs are accented.

Test II. This test is administered by handing the picture booklet to the child with the following instructions:

> When I say "go," call aloud the names of these pictures as quickly and as clearly as you can. Start at the top of the page and read across as you would in a book. When you finish a page, turn it over and read the next one. Do this until you have read them all. Do you understand? Remember, read them as quickly and as clearly as possible. Go.

This test is scored in the same manner as is Test I. The examiner has in front of him a scoring sheet containing the names of all of the pictures in the booklet on which the same symbols may be used to indicate the errors made. The total time taken to name all of the pictures in the test proper is recorded, as well as the total number of errors.

TABLE 1

MEAN SCORES AND SIGMAS FOR EACH GRADE

Fifty cases in each grade. Time scores in seconds.

Grade	Reading speech test				Picture articulation test				Spontaneous speech test				
	Mean time	σ time	Mean errors	σ errors	Mean time	σ time	Mean errors	σ errors	Mean time	σ time	Mean words	σ words	Rate per sec.
I	376	153	49.6	23	222	56	3.7	4.2	45	17.4	56.6	20.6	1.26
II	186	133	22.0	23	230	60	11.2	9	49	22	58.2	27	1.19
III	89.7	35	9.8	6.7	161	46	5.5	4	44.2	19.9	59.9	20	1.35
IV	65	30	5.3	6.5	107	38	2.7	1.4	38.2	15.9	73.5	32.5	1.92
V	54.4	11.9	5.0	4.0	91.6	21	2.7	1.4	38.8	15.9	68.4	21.5	1.76
VI	45.2	10.6	4.8	2.8	76.6	16.7	1.5	3.5	42.2	19.4	67.4	27	1.60
VII	43	12.5	5.0	4.0	70	44	.8	.3	38.2	13.9	59.2	20.7	1.55
VIII	44.2	5.9	5.0	4.5	66.6	10.2	.7	.3	36	17	71.6	25	1.99

Test III. In the administration of Test III the examiner simply asked the questions in order and timed the period of speaking with a stop-watch held in one hand and counted the number of words with a hand tally held in the other hand. The total time for the answer to each question and the total number of words used in response to each question are recorded. The speech rate is computed from these data.

Tentative Norms of Test Performance

In order to establish tentative norms of test performance the above-described tests were administered to fifty subjects in each grade from first to eighth inclusive. Table 1 shows the results obtained.

Intertest and Intelligence Correlations

Each test in the battery was correlated with each of the other tests and with intelligence quotients. For this purpose Grades VI, VII, and VIII of the control groups were lumped together. Since the differences in mean scores made by these groups were relatively slight, it was considered that the combined group used in this part of the study was relatively homogeneous. Table 2 shows the results obtained.

TABLE 2

Intertest and Intelligence Correlations

Items correlated	*N*	*r*	*P.E.*(*r*)
Reading speech time—IQ	150	—.055	.058
Picture-naming time—IQ	150	—.170	.056
Spontaneous speech rate—IQ	150	.168	.056
Picture-naming time—spontaneous speech rate	150	—.364	.050
Reading speech time—spontaneous speech rate	150	—.058	.058
Picture-naming time—reading speech rate	150	.510	.043

Reliability of the Tests

As a means of checking the reliability of the tests they were readministered to a group of fifty of the control subjects from the fifth and sixth grades about two months after the first administration. This group was chosen primarily because of its availability,

but also because it represented the middle range of those tested and because the group was fairly homogeneous. It was thought that the dual-administration method might be partially responsible for the high reliabilities obtained and for comparative purposes the first half of the reading speech test was correlated against the second half, correcting by the Spearman-Brown formula.

A second administration of the tests was also given to fifty cases from the middle range of the stuttering group. Although this group was not as homogeneous as the control group it was thought advisable to get some indication of how the tests hung together with the speech defectives. Table 3 shows the results obtained.

TABLE 3
RELIABILITY INDICES

	Normals			Stutterers		
	N	*r*	*P.E.*(*r*)	*N*	*r*	*P.E.*(*r*)
Reading speech time, two administrations	50	.94	.01	50	.87	.02
Picture-naming time, two administrations	50	.92	.02	50	.98	.01
Spontaneous speech rate, two administrations	50	.73	.05	50	.98	.01
Reading speech time, 1st half *vs.* 2nd half	50	.87	.03			

VALIDITY OF THE TESTS

The determination of the validity of these tests as measures of speech ability, or the lack of it, encountered the usual difficulty of determining adequate criteria with which to correlate test results. Teachers, acting as judges, are quite apt to confuse speech ability, which in this case is intended to mean the mere functioning of vocal apparatus, with the use of language, vocabulary size, and similar factors.

The first step in checking the validity was to administer the tests to as many stutterers as were available in order to see how this group compared, grade for grade, with the control group. While the number of cases in each grade is too small to be of much significance, the results are at least indicative of the trend. Table 4 shows the results obtained by giving the tests to a group of stutterers. It is to be compared with Table 1 which represents the same material for the control group.

TABLE 4
MEAN SCORES AND SIGMAS OF STUTTERERS

	Reading speech test				Picture articulation test				Spontaneous speech test				
N	Mean time	σ time	Mean errors	σ errors	Mean time	σ time	Mean errors	σ errors	Mean time	σ time	Mean words	σ words	Rate per sec.
14	200	90	35.7	19	214.5	71	12.9	9.9	62.5	27.5	74.3	22.8	1.19
16	128	71.5	21.3	8.2	190.6	59.8	10.6	8.3	46.8	15	76.8	21.5	1.64
19	88.8	33.8	18.1	13.3	140	55	10.2	13.9	67.8	30.8	104.5	47.3	1.54
19	72.8	21.4	15.5	12.7	125	55	13.9	18.8	61.5	50.5	101.8	57.5	1.65
13	72	23	15.7	8.2	106.8	32.5	4.4	3.7	60.8	24.8	100.8	34.8	1.66
11	75	52	19.5	16.7	107.8	31.8	8.4	10.2	46.8	20.3	80.5	32.5	1.72
7	59.2	17.6	14.6	10.6	86.4	20.2	5.4	4.5	69.5	39.5	117.5	42.5	1.69

TABLE 5
PERCENTAGES OF CONTROL GROUPS EXCEEDING MEANS OF STUTTERERS

Grade	*N*	% controls equal or exceeding mean time of stutterers in reading speech test	% controls equal or exceeding mean errors of stutterers in reading speech test	% controls equal or exceeding mean time of stutterers in picture-naming test	% controls equal or exceeding mean errors of stutterers in picture-naming test	% controls equal or exceeding mean of stutterers:	
						in no. of words used	in rate of speech
II	14	70	82	38	66	20	62
III	16	82	56	76	80	16	28
IV	19	90	96	84	100	8	78
V	19	94	96	90	100	6	64
VI	13	96	100	96	94	6	54
VII	11	98	94	84	96	20	36
VIII	7	100	96	90	98	4	90

Using Tables 1 and 4, the next step was to compute the percentage of the control group who equalled or exceeded the mean scores of the stuttering group. It would be expected that, if the tests had any degree of validity, this percentage would be well over 50. On the other hand, it should be pointed out that, although the stuttering group was highly selected, the control groups were unselected and contained, in all probability, a fair proportion of subjects with poor speech ability, which would tend to make the percentages fall short of 100. Table 5 shows the percentages of the control who equaled or exceeded the mean scores of the stuttering group. In the reading speech test and in the picture naming test, time and errors were considered. In the spontaneous speech test, both the rate of speech and the total number of words used in response to the three test questions were considered.

It will be seen that except for an inversion in the picture-naming test time in the second grade, the reading speech and picture-naming tests are fairly differentiating, especially above the third grade. Speech rate, as will be pointed out shortly, is of little or no value as an index of speech difficulty.

Since the first two tests of the battery seemed to be fairly differentiating, the next step was to correlate the test results with the ratings of teachers of the ability in question. Grades V and VI were used in this part of the study as well as a group of stutterers, ranging from Grades III to VI, who happened to be under one teacher. These correlations are shown in Table 6.

TABLE 6

CORRELATIONS BETWEEN TEACHER'S RATINGS AND TEST SCORES

Teacher's rating of speech ability with:	Group	*N*	*r*	*P.E.*(*r*)
Reading speech time	5th-gr. control	50	.31	.09
Reading speech time	6th-gr. control	50	.52	.07
Reading speech time	stutterers 3-6 gr.	50	.26	.09
Picture-naming time	5th-gr. control	50	.23	.09
Picture-naming time	6th-gr. control	50	.19	.09
Picture-naming time	stutterers 3-6 gr.	50	.28	.09

Speech Rate of Stutterers

It was naïvely supposed at the outset of this study that an impediment in speech was a factor which would impede, among other things, the rate of speech. The fact that Stinchfield included speech rate in her series seemed to justify this assumption. That this was not the case was soon apparent, and, as may be seen in Table 5, the mean speech rate of the stutterers was exceeded by the mean rate of the control group in only four out of the seven grades dealt with in this study. In order to determine whether or not any difference prevailed between the stuttering and the control groups, Grades IV to VIII inclusive in both groups were combined and the reliability of the difference determined. As may be seen in Tables 1 and 4, these grades may be combined and still be a homogeneous group as far as rate of speech is concerned. Table 7 shows the mean speech rate, the sigma of the mean, the difference between the means, and the sigma of the difference.

TABLE 7

Difference in Spontaneous Speech Rate between Stutterers and Non-Stutterers

	Grades	N	Mean rate of speech	Mean
Control	4-8 inc.	250	1.764	.097
Stutterers	4-8 inc.	68	1.652	1.56
		Difference	.112	
		$\sigma_{diff.}$	1.56	
		$\frac{\text{Diff.}}{\sigma_{diff.}}$	.0718	

Total Words Used by Stutterers

Referring again to Table 5, it is seen that the total number of words used in response to the three standard questions by the stuttering group was without exception far in excess of the total number of words used by the control group. Here again there was not a significant increase in the number of words used after the third grade in either the stuttering or the control group, and again Grades IV to VIII were combined in each case in order to determine the difference in the number of words used. This difference is shown in Table 8.

TABLE 8

DIFFERENCE BETWEEN STUTTERERS AND CONTROLS IN MEAN NUMBER OF WORDS USED

	Grades	N	Mean number of words used in answer to three questions	σ Mean
Control	4-8 inc.	250	68.02	1.61
Stutterers	4-8 inc.	68	101.02	5.20
		Difference	33.00	
		$\sigma_{diff.}$	5.44	
		$\frac{\text{Diff.}}{\sigma_{diff.}}$	6.07	

CONCLUSIONS

1. Of the three tests originally chosen, two seem to be fairly good quantitative measures of speech. These tests are oral reading rate and rate of picture naming (Table 5).

2. These two tests are also so designed as to be qualitative measures of speech difficulty.

3. That these two tests are measuring something other than intelligence is indicated by their low correlations with intelligence (Table 2).

4. These tests have a positive, though not very high, correlation (av. .30) with teachers' ratings of speech ability (Table 6).

5. Although these tests were made as simple as possible in order to measure a minimum of reading ability, there is a gradual improvement in the reading test from the first grades through the fifth grades of elementary school, and a gradual improvement in the picture tests through the eighth grades. Hence grade norms are considered essential (Table 1).

6. The tests as used proved to be sufficiently reliable for practical purposes (Table 3).

7. Within the scope of this study, there were no significant differences between stutterers and normals in the rate of spontaneous speech, and such a test, therefore, was not considered to be a measure of speech ability (Tables 5 and 7).

8. Within the scope of this experiment, there is no constant or significant increase in speech rate from the first through the eighth

grades of elementary school in either normal or stuttering groups (Tables 1 and 4).

9. In response to three standard questions, the stutterers used 50% more words than the non-stutterers. This difference is 6.07 times the sigma of the difference (Table 8).

References

1. Baker, —, & Baker, —. The Bobbs-Merrill primer. Indianapolis: Bobbs-Merrill, 1929.
2. ————. The Bobbs-Merrill first reader. Indianapolis: Bobbs-Merrill, 1929.
3. Blanton, —, & Stinchfield, S. M. Speech measurements: a series of graded speech tests. Chicago: Stoelting, 1924.
4. Stinchfield, S. M. The formulation and standardization of a series of graded speech tests. *Psychol. Monog.*, 1924, **33**, No. 149. Pp. 54.
5. ————. The psychology of speech. Boston: Expression Co., 1928. Pp. ix+331.
6. Stoddard, C. Articulation tests. Detroit, Mich.: Detroit City Schools, 1924.

Washington University
St. Louis, Missouri

UN ESSAI FAIT POUR MESURER DE CERTAINES PHASES DE LA PAROLE

(Résumé)

Dans un essai fait pour mesurer quantitativement la capacité de s'exprimer par la parole, on a employé trois mesures: (1) la durée de la lecture d'un passage simple, (2) le temps mis à nommer une série de tableaux, et (3) la vitesse de la parole spontanée. D'entre celles-ci, seulement les deux premières se sont montrées capables de distinguer entre les bègues et les non-bègues. La constance du test a été d'environ 0,90. On a déterminé la validité (1) en déterminant le pourcentage par lequel les groupes de contrôles, égalés selon l'année scolaire aux groupes de bègues, ont dépassé le résultat moyen des bègues, (2) au moyen des corrélations peu élevées entre les résultats des tests et l'intelligence et (3) au moyen de la corrélation entre les résultats des tests et les évaluations de la capacité de s'exprimer par la parole faites par les maîtres d'école.

Les progrès graduels faits dans la dureé de la lecture, même avec des matières simples, et dans le temps mis à nommer des tableaux indique la nécessité de normes adéquates d'âge ou d'année scolaire. On donne des normes tentatives d'âge dans cette étude. La vitesse de la parole s'est montrée de peu de valeur ou d'aucune valeur pour distinguer entre les bègues et le non-bègues. La vitesse de la parole s'est montrée assez constante dans les huit premières années scolaires chez les groupes des bègues et les groupes des non-bègues.

Quand l'on a comparé les groupes des bègues et ceux des non-bègues sur la base du nombre de mots employés dans la réponse à trois questions simples, on a constaté que les bègues ont employé 50% plus de mots que les normaux.

HAWTHORNE

EIN VERSUCH, GEWISSE BESTANDTEILE DER SPRACHE ZU MESSEN

(Referat)

In einem Versuch, die Sprachfähigkeit quantitativ zu messen, wurden drei Massstabe angewendet: (1) die zum Durchlesen einer einfachen Stelle [passage] nötige Zeit [reading time]; (2) die zur Bennennung der Mitglieder einer Bilderserie nötige Zeit; und (3) die Schnelligkeit der spontanen Sprache. Es zeigte sich, dass nur durch die ersten zwei dieser Massstabe zwischen Stotterern und Nicht-Stotterern unterschieden werden konnte. Die Zuverlässigkeitszahl [reliability] der Prüfungen betrug ungefähr .90. Die Validität wurde bestimmt: (1) dadurch, dass man den prozentmässigen Unterschied zu Günsten der Kontrollgruppen bestimmte, zwischen den durchschnittlichen Leistungen [mean score] der Stotterern einerseits und den Leistungen von Kontrollgruppen, die in Bezug auf Schulklasse den Stotterergruppen parallel standen [equated by grade placement] anderseits; (2) durch die niedrigen Korrelationen zwischen den Testleistungen [test scores] und der Intelligenz; und (3) durch die Korrelation zwischen den Testleistungen einerseits und den von Lehrern gemachten Abschätzungen [teachers' ratings] der Sprachfähigkeit der Vpp. anderseits. Die allmähliche Beschleunigung (Besserung) der Lesezeit [improvement in reading time], auch an einfachem Lesestoff, und der zur Benennung der Bilder nötigen Zeit [picture-naming time] weist auf die Notwendigkeit angemessener Bestimmungen der Normalleistungen nach Alter oder Schulklasse hin [necessity of adequate age or grade norms]. Es werden in der vorliegenden Arbeit vorläufige Bestimmungen dieser Art, nach Schulklasse, angegeben. Der Wert der Bestimmung der Schnelligkeit der Sprache erwies sich bei der Differenzierung zwischen Stotterern und Nicht-Stotterern als gering. Es zeigte sich, dass die Schnelligkeit der Sprache sowohl bei den Stotterern wie bei den Nicht-Stotterern sich von der ersten bis durch die achte Schulklasse ziemlich konstant erhielt. Man verglich die Stotterergruppen mit den Nicht-Stotterergruppen auch in Bezug auf die Zahl der bei der Beantwortung von drei einfachen Fragen benutzten Wörter. Es zeigte sich, dass die Stotterer hierzu 50% mehr Wörter gebrauchten, als die Nicht-Stotterer.

HAWTHORNE

PSYCHOPHYSIOLOGICAL MEASUREMENTS DURING SOMNOLENCE INDUCED BY PHENOBARBITAL (LUMINAL)*

From the Departments of Psychiatry and Psychology, Institute of Human Relations, Yale University

LOUIS H. COHEN, ROBERT R. SEARS,[1] STANLEY B. LINDLEY, WALTER C. SHIPLEY, AND PAULINE K. SNEDDEN

Quantitative measures of certain physiologic and psychologic activities have been obtained during the somnolent stage of phenobarbital (luminal) intoxication in a patient addicted to episodic usage of the drug.

CASE HISTORY

Mrs. J. R., a white Jewish woman, age 45, was admitted to the New Haven Hospital, June 24, 1931, in a stuporous state. She reported that she could see things only as shadows and that she felt very sleepy. With the exception of constricted pupils, stertorous breathing, slurring speech, and slight edema of the ankles, physical examination was essentially negative. In 3 days she had completely recovered from her stupor and appeared to be a rather active, intelligent woman in whom no particular mental or physical difficulties were noticeable except for an amputation type of anesthesia and analgesia of the left arm with astereognosis and muscular weakness. There was no free HCl in the gastric juice; laboratory findings were otherwise negative.

Her family history was essentially negative. She had the usual childhood illnesses and several abdominal operations. After one operation, at 37, her arms and legs became "paralyzed" and it was necessary to remain in bed for several months. At this time also it was noted that on an occasion when it became necessary to sign some papers she suddenly became blind. It required 2 months for the patient to regain her sight and the use of her limbs. While

*Accepted for publication by Walter Miles of the Editorial Board and received in the Editorial Office, June 19, 1933.

[1]Blossom Fellow in Neuro-Anatomy, 1931-1932.

convalescing from a later abdominal operation she again developed a quadriparesis of the same nature as before, and which, with the exception of the left arm, recovered completely in a few months. For the last several years generalized edema had come and gone with dramatic suddenness; the patient on one occasion gained 16 lbs. in 36 hours, and lost this weight over about 2 weeks; altogether she reported 8 of these attacks.

In April, 1931, the patient had the first of her somnolent attacks which lasted 2 or 3 days. Since that time these attacks had become more frequent and of longer duration. She reported that an attack tended to come on slowly, that she would become very frightened and would attempt to fight it off but that invariably she would succumb. The somnolent period would last from 3 to 8 days with apparent gradual lessening of the stupor; on recovery the patient would have severe headache, some blurring of vision, and difficulty with her speech.

The patient was observed during numerous attacks of somnolence during her stay in the Psychiatric Clinic. These recurred irregularly every 2 or 3 weeks and lasted from 3 to 7 days. The physiological sequence of events was almost invariably as follows: The patient would begin to gain weight, and edema would appear in her legs, arms, and occasionally in her abdomen, chest, and face; after 4 to 10 lbs. were gained, in the course of 3 to 14 days, diarrhea of 5 to 12 stools per day would set in (otherwise she was constipated or had 1 or 2 bowel movements per day) and after 2 to 5 days of diarrhea she would go rather swiftly into a somnolent state. The transition from complete lucidity to somnolence would occupy a few hours or would be reported by the nurse as having suddenly occurred. During the somnolent state skin temperature variations were quite different from those during the lucid state (2).

It was eventually discovered that the patient was a phenobarbital (luminal) addict but at the time of the experiments reported below this was not known. For this reason the amount of drug which had been taken while these studies were made cannot be reported. That very large doses were probably taken may be surmised since it was later demonstrated that the patient had an extremely high tolerance for the drug. When isolated from the drug for two months no edema, diarrhea, or somnolence occurred; when allowed to procure the drug the cycle would recur.

The diagnostic impression was that the patient was a psychopathic personality of the hysterical type with episodic drug (luminal) addiction.

I. Patellar Tendon Reflex

It has been shown by Lee and Kleitman (3), Tuttle (9), Bass (1), and others that the excitability of the knee-jerk as measured by its amplitude is either markedly diminished or altogether abolished during sleep. Wilson (10) adds, further, that it disappears entirely during attacks of idiopathic narcolepsy. There was no diminution in its amplitude during hypnosis, according to Bass. Inasmuch as the somnolent state of luminal intoxication resembles, at least superficially, all three of these conditions, particularly sleep, it seemed desirable to determine the relative excitability of the knee-jerk in this patient.

Fortunately the apparatus described by Bass (1) was available. The patient, resting in a modified barber's chair, lay in a horizontal reclining position with both legs hanging free from the knee. Stimuli were delivered by the falling of a 300-gram pendulum through an arc of 45° upon a wire-wound bamboo rod placed tightly against the patellar tendon of the subject's left leg. Responses were recorded on stylograph paper by a lever operated by a thread glued to the patient's leg. The occurrence of stimuli was automatically recorded by an electromagnetic marker, the circuit through which was closed by the incidence of the pendulum's blow against the rod. In order that the millimeter-measurements of the amplitudes of the recorded reflexes might be strictly comparable from day to day, care was taken to keep apparatus conditions as constant as possible. The spot at which the thread was fixed to the leg was inked at the first session and kept constant thereafter.

Measurements were made with two types of stimulational setting: (*a*) with and (*b*) without a secondary facilitating stimulus. In (*a*) the subject was instructed to lie as quietly as possible. Blows to the tendon were presented at regular 10-second intervals unaccompanied by any other controlled stimulation. In (*b*), the experimental conditions and tendon stimulation were the same except that a buzzer was sounded every 60 seconds and the patient instructed to press a telegraph key under her right index finger in response to it. At no time did this buzzer stimulus and reaction

coincide with a tendon stimulus, since it occurred 5 seconds after each sixth consecutive blow to the tendon.

In Table 1 are shown the comparative mean amplitudes under each of three conditions (luminal somnolence, "very drowsy," and normal waking) with unfacilitated patellar stimulation. The somnolence test was made on August 9, 1931, during a period characterized by marked stupor. On August 12 the patient was able to sit up and read comfortably; and, although all evidence of intoxication had passed off with the exception of a feeling of "pressure" on the temporal aspects of the head, when placed in the apparatus, the patient complained of being very drowsy. It will be

TABLE 1
MEAN AMPLITUDE (IN MM.) OF THE FIRST CONSECUTIVE 100 PATELLAR REFLEXES ELICITED ON EACH OF 3 OCCASIONS

Date	Condition	Stimulation: Blow to tendon	Stimulation: Blow to tendon; buzzer at 60 secs.
8/9/31	Somnolence	9.23± .46	39.06±1.19
8/12/31	Very drowsy	4.53± .53	
8/13/31	Waking	35.84±1.09	36.57± .96

noted that the reflex amplitudes were reliably less than those obtained during somnolence. During a second series of stimulations on the same occasion she was wide awake at first and then went to sleep for a few minutes. The mean amplitudes of the waking half of this series (29.2 mm.) were similar to those obtained the following day (August 13) when the "waking" control series was given, while those obtained during the second half of the series, in sleep, were reduced to an average of 1.5 mm.[2]

When the patient was forced to react to the buzzer every 60 seconds the excitability of the reflex was as great during somnolence as during a normal control (waking) period. This secondary facilitating stimulus and reaction apparently nullified the depressant effects of the drug.

Summary. The excitability of the patellar tendon reflex as measured by comparative amplitudes was greater during the period of

[2]The values obtained in this extra series are not strictly comparable with the other values reported, since negative adaptation during the first 100 reactions would influence these later amplitudes to an unknown extent. They have not, therefore, been included in Table 1.

somnolence than during drowsiness but was markedly less than during a normal waking period. Addition of a secondary facilitating stimulus-reaction pattern (pressing key to buzzer) served to raise the somnolence excitability to the waking norm.

II. Conditioned Finger Retraction

While it is obvious that modification of behavior takes place in sleep as well as in other conditions, such as hypnosis, Pavlov (5) has demonstrated the extreme difficulty or impossibility of establishing overt conditioned reactions during sleep. Scott (6) has shown that it is possible to establish them during hypnosis. The efficiency of the conditioning process, then, provides another point at which a comparison between luminal somnolence and "normal" sleep can be made.

The patient was placed in a Morris chair, her right index finger resting on one electrode of a faradic circuit. The other electrode (diffuse) was strapped to the wrist. Movements of the finger were recorded on smoked paper by a sensitive ammeter the activity of which was controlled by changes of pressure in a telephone-receiver diaphragm which supported the finger electrode. The sound of an electric buzzer served as a conditioned stimulus and an electric shock as the unconditioned or adequate stimulus.

The procedure remained the same for the first five days of the experiment. The buzzer was sounded alone 3 times in succession at intervals of 38 seconds. These stimuli were followed by a series of 50 reinforcements in which the buzzer preceded the shock by approximately ½ second, the interval between reinforcements being 38 seconds. The buzzer was again sounded alone to determine the presence or absence of conditioning. Fifty further reinforcements were given followed by 5 test trials (buzzer alone). On the sixth day the procedure was the same except that each of the two groups of reinforcements was reduced from 50 to 25, and 2 test trials were given between them while at the end only 4 test trials were given.

From Table 2 it may be seen that no evidence of conditioning was obtained during the first five days, i.e., with 500 reinforcements. During this period the patient was in an extremely somnolent condition and reported hearing the buzzer only when she was aroused by the experimenter. During the conditioning process she was not, of course, aroused or disturbed in any way. The unconditioned re-

TABLE 2
SCHEMA OF THE CONDITIONED-RESPONSE EXPERIMENT

Date	Hour	Condition	Reinforcements	Test trials	Conditioned responses
10/16/31	3:30 P.M.	Somnolent	100	9	0
10/17/31	11:15 A.M.	Somnolent	100	9	0
10/18/31	10:15 A.M.	Somnolent	100	9	0
10/19/31	9:30 A.M.	Somnolent	100	9	0
10/20/31	9:30 A.M.	Somnolent	100	9	0
10/21/31	9:30 A.M.	Waking	50	9	3+; 4—

actions, however, appeared normal, the patient frequently returning her finger to the electrode without assistance.

On the sixth day the somnolence disappeared and the patient seemed to be quite wakeful. The 3 initial buzzer tests elicited negative reactions, that is, pressure on the electrode rather than retraction from it. The first test after 25 reinforcements elicited a strong positive reaction but the second produced another negative one. After another 25 reinforcements 4 more buzzer tests were given and she responded positively to the first and third but negatively (pressure) to the second and fourth.

There is little doubt that the conditioning took place largely on the last day, although the presence of the negative reactions at the beginning of the session indicates that some modification must have taken place during the preceding 5 sessions. Inasmuch as no reaction of any kind was obtained until the last day, motor learning cannot be said to have occurred overtly until that time. It does not seem probable that 500 reinforcements were required for conditioning, since a response of the same nature but to a tactual stimulus was established in this patient several weeks previously in less than 100 trials (Sears and Cohen, 7). Other investigators also have shown the ease with which conditioned finger retraction is ordinarily established (Scott, 6; Wolfle, 11; and Shipley, 8).

Summary. It was found impossible to establish a conditioned finger retraction to a buzzer during the somnolent period, but some implicit learning apparently took place. The response was readily established in the waking state.

III. MAINTENANCE OF ATTENTION

In order to obtain some objective quantitative measure of the patient's attentional capacity at a relatively simple level of behavior a test of efficiency in subvocal counting of auditory stimuli was

employed. The patient was placed in the modified barber's chair described in the account of the patellar-reflex experiment, her right index finger resting on a bell-push at her side. A buzzer, regulated by an electric clock, was sounded for ½ second at 5-second intervals. Incidence of stimuli and responses was recorded electromagnetically on a polygraph.

Two types of subvocal counting were used. In the first, the patient was requested to press the button at regular intervals of 5, 10, 15, 20, 25, and 35 seconds. That is, after the first response, a certain number of buzzer sounds (depending on which of the above intervals was being used) was allowed to pass without the reaction of pressing the button. The instructions were:

> You will hear the buzzer at regular intervals (or, at the same interval as before), and I want you to press the button every time (or, every second, third, fourth, fifth, or seventh time) you hear the buzzer.

The second type of subvocal counting was more complex. It involved a *pattern* of intervals between reactions rather than a single repeated interval. The instructions were:

> You will hear the buzzer at the same intervals as before, but now I want you to press the button every second and every third time alternately.

Seven series of this type were used, each in turn having one more variable in its pattern. Since memory ability for digits would be important with such complicated patterns, the instructions were always repeated as many times as necessary for accurate memorization and the patient was asked to repeat the series of numbers again after completing each experiment.

Twenty-five reactions were obtained with each of the 13 series of patterns in both the normal and somnolent condition. In scoring the results, any failure to respond at the proper time was counted an error no matter what the specific nature of the incorrect response might be, i.e., anticipation, omission, or procrastination.

In Table 3 are shown the percentages of correct responses in the various series under the two conditions. During somnolence the efficiency of the measured process was decidedly lower than during the normal session, although the latter preceded the former by three days; any practice effect should have tended to lessen such a

difference as is shown. The difference between the mean percentages of correct responses is 9.85%, with a P.E. of 2.72; the critical ratio is 3.62. A coefficient of reliability obtained by correlating the two columns of values in Table 3 is —.52 in spite of the difference in the patient's condition at the two testings.

The data in the column headed "% difference from normal" present several striking differences between the somnolent and normal state data. If the Type-I data (Series 1-6) are considered alone it may be seen that the difference between the somnolent and normal state is always a negative one which for the six series averages —18.66%, and, if Series 1 is omitted, averages —22.4%. In the Type-II data (Series 7-13), on the other hand, the differences are unexpectedly smaller; of the seven series, five show minus differences which average —7.2% and two show plus differences averaging +10%. The significance of these figures lies not so much in the quantitative difference as in the fact that apparently the decrease of efficiency in maintenance of attention was markedly less for the more complex task. Whereas the total number of errors was greater during the somnolent series than during the normal, these errors were predominantly more numerous in the Type-I (simple) series than in Type II (complex).

TABLE 3

EFFICIENCY IN SUBVOCAL COUNTING DURING SOMNOLENT AND NORMAL WAKING CONDITIONS

Values given are percentages of correct reactions recorded in the 25 actual reactions of each series.

Series	Pattern	Normal	Somnolent	% different from normal
1	5″	100	100	0
2	10″	96	88	— 8
3	15″	100	68	—32
4	20″	100	80	—20
5	25″	96	80	—16
6	35	100	64	—36
7	10″-15″	96	88	— 8
8	10″-20″-35″	92	88	— 4
9	10-15-5-25	100	96	— 4
10	10-5-15-15-5	100	88	—12
11	10-20-10-10-15-25	88	92	+ 4
12	15-25-10-5-20-15-5	100	92	— 8
13	5-15-20-5-10-25-15-10	64	80	+16
	Mean	94.77	84.92	
	$P.E._M$	1.98	1.87	
	$P.E._{diff.}$	2.72		
	Critical ratio	3.62		

These subvocal counting tasks are of the utmost simplicity from the sensory and motor standpoints. Although reaction-time to an auditory stimulus was decreased in the somnolent state, speed could not be considered a factor in the measurement of efficiency on this task. Likewise, the patient's perfect score on the first series during somnolence indicated a sufficient capacity to react to the stimulus. It may be assumed, therefore, that whatever disturbance appeared in this experiment was a representation of disturbed function at the higher associative levels, namely, of the attentional processes.

Summary. The ability to maintain attention was decreased during the somnolent period, but the chief reduction was with a relatively simple and monotonous task, while relatively less decrement was observed in the efficiency of responses to a complex task.

IV. Auditory Reaction-Time

A comparison was made of reaction-times to auditory stimuli obtained under waking and somnolent conditions. The subject sat comfortably at a table, her right hand resting on a telegraph key the total closure-excursion of which was 2 millimeters. The weight required for closure was about 150 grams. Titchener's abbreviated motor instructions were read aloud:

> When you hear the click from the sounder, press down on the key. Press down as quickly as possible.

An offset from the relay (stimulus) and an electromagnetic marker activated by the telegraph key (response) were placed before the slit of a motor-driven continuous-roll camera. A beam of light directed toward the slit recorded the shadows of these markers on the sensitive paper. A second beam of light, interrupted at a constant rate of 50 times per second, was focussed on the camera slit also and provided a time line which allowed accurate measurement, by interpolation, to within 5 sigma of the actual value of any recorded interval.

The patient was given 100 stimulations per day, 50 for each hand, for 7 days (August 29—September 4, 1931). These days included the last 5 of a 6-day somnolent period and the first 2 of the following waking period. A nurse was present at all times to be certain the patient was as attentive as her condition would permit. In Table 4 are shown the means and sigmas of the distribution on each of the

7 days. At the foot of the table is shown the range obtained by Miles (4) for auditory reaction-times from a group of 100 unselected adults using their preferred hands (ages 20-80).

TABLE 4

MEANS AND STANDARD DEVIATIONS OF REACTION-TIME LATENCIES FOR EACH HAND ON EACH OF 5 SOMNOLENT DAYS (8/29-9/2) AND THE FOLLOWING 2 WAKING DAYS (9/3, 9/4), TOGETHER WITH THE CRITICAL RATIO OF THE DIFFERENCES BETWEEN THE SEVERAL MEANS

	Means		D	Standard deviations	
Date	Right	Left	P.E.$_{diff.}$	Right	Left
8/29	585.8	701.4	2.21	82.26	527.67
8/30	591.5	664.0	2.88	131.66	185.68
8/31	681.2	685.8	1.58	179.87	243.39
9/1	539.5	591.8	3.75	99.77	111.15
9/2	352.8	345.9	.57	64.68	106.71
9/3	293.4	275.8	1.67	74.08	81.05
9/4	203.5	203.5	.00	53.49	65.99
Miles: range for 100 unselected adults				100-817	

It will be noted that the latency during somnolence was materially greater than during the 2 waking days at the end, but that the last 3 days show a progressive decrease.

Although the patient believed her left hand to be partially paralyzed, no objective evidence of such dysfunction is found in the reaction-time data. The first sets of measurements show some difference between the two hands but this is not highly reliable; the later figures show almost no difference. Inasmuch as the left hand was rarely used, it apparently had become somewhat weak; the practice of the first 4 days seemed to have brought its speed up to that of the right. She remarked spontaneously after the last session that she supposed the left hand was slower, but made no comment when informed that it was not.

Summary. During somnolence the auditory reaction-time was increased to nearly three times its normal value as measured immediately after the somnolent period. No significant difference was found between the reaction-times of the two hands after four days of practice, although the patient believed the left hand to be paralyzed.

V. Voluntary Reactions to a Pattern of Sounds

The patient's voluntary reactions to a relatively simple rhythmic sound pattern were recorded for 12 successive days during which time she went from an entirely lucid state into somnolence and then returned to the lucid condition.

A pattern composed of four sounds of approximately equal intensity, duration, and frequency was selected. The pattern is of this form 1--1-1-1--, i.e., three beats separated by equal intervals and the fourth beat separated from its neighbor by twice the interval between the other beats. The rate of presentation was one cycle of the rhythmic pattern, four stimuli, per one second. The duration of the presentation of the stimuli was 1½ minutes, during which time the pattern was repeated 90 times. These sounds were presented to the subject through the medium of a nearby electromagnetic relay in the form of metallic clicks. The subject reacted by pressing a key at the incidence of each stimulus. Photographic registration was employed. The instructions read to the patient were as follows:

> You are to press or tap the key in front of you so that the click at the end of your stroke will coincide with the click from the sounder. Pay close attention to the clicks from the sounder in order to be sure that the clicks made by you occur at exactly the same instant. I will give a warning signal before the clicks begin. I will say "ready" about three seconds before the clicks begin.

The patient had participated in a similar experiment a week before beginning the experiment under discussion. The data from the two experiments were in agreement where the conditions were comparable.

The patient's reactions to the sound pattern were scored on the basis of the number of sigma by which each reaction occurred antecedent to, or lagged behind, the "stimulus." The error was measured to 1/200th of a second. The reactions that occurred before the "stimuli" were considered anticipatory reactions. Those that occurred after the incidence of stimulation were considered latent reactions. Actually, all responses less than about 150 sigma [Woodrow's (12) range is 129-180] after the instant of stimulation might well be described as anticipatory. The formula for the

calculation of the constant error from the combination of the three distributions was as follows.

$$\text{C.E.} = \frac{\text{(anticipatory frequency x deviation)} - \text{(latent fd.)}}{N \text{ of antic.} + N \text{ of latent} + N \text{ of coincident reactions}}$$

This experiment began on August 24, 1931, and was continued, being given at the same time each day, until September 4, 1931. The patient lapsed into the somnolent state during the period of experimentation on the morning of August 29 and did not show obvious improvement in her condition until September 2. September 3 and 4 were pronounced days of complete lucidity for the subject. Hence the experiment deals with six normal days, four preceding, and two following a somnolent attack. Included, also, are the data for the day of the onset of the attack, August 28, and for the first partially lucid day following the somnolent period, September 2. During the days August 29 to September 1, inclusive, the experiments were carried on as usual but the patient's responses were of such a different nature from those of the other days that they do not permit of similar treatment, at least for the present. The patient's reactions were evenly and relatively widely spaced pressures upon the reaction key. No grouping of responses was observable and they were separated on the average by about 400 sigma.

Reference to Table 5 clearly reveals the trend of the patient's behavior throughout the series. For the first day the constant error is —14.589 sigma. On day 8/26 the constant error is reduced

TABLE 5

CONSTANT ERRORS (IN SIGMA UNITS) OF THE REACTIONS TO THE RHYTHMIC SOUND PATTERN

— signifies that constant error was in direction of latency of reaction.

Date	Constant error	S.D.
8/24	— 14.598±1.852	35.767
8/25	— 15.348±1.568	30.315
8/26	— 11.546±1.622	31.661
8/27	— 26.008±1.511	29.266
8/28	—156.628±6.221	115.884
8/29—9/1*		
9/2	—186.118±8.399	152.802
9/3	— 8.944±2.124	41.429
9/4	— 1.625±1.318	25.11

*Note: from 8/29 to 9/1 inclusive the responses were of such a different nature from those dealt with above that they could not be treated in the same way as the other data.

in latency sufficiently to allow one to hazard that learning is occurring. But on 8/27 the constant error suddenly reverts back to a markedly increased latency, —26.008 sigma, and on the next day the patient was in one of her somnolent states. Fortunately, on day 8/28 the experiment caught the patient going rapidly into a somnolent state. When she appeared in the laboratory on that morning she showed a little more than her usual ataxia and in addition complained of slight dizziness. The experiment consumed about 3 or 4 minutes, and, though the patient carried through the entire 1½ minutes of work, at the end of that time she was unable to rise from the chair unaided. She was immediately returned to her room and in a short time was pronounced to be in a somnolent state. On each of the next four days the patient, in charge of a nurse, was put through the same experimental regime, except that a series of reaction-time measurements completed each day's work. During these days the subject could be made to understand the instructions and carried out her reactions as well as she could. She always completed the 1½ minutes of work, but her movements were slow and ungrouped, showing no configuration comparable to the stimulus pattern. The movements were ponderous and seemed to be accomplished with a great expenditure of energy and with considerable involvement of the other musculature of the body. For instance, she would raise and lower her head in time with the movements of the hand with which she was tapping. Once the sound pattern had started, the patient did not lapse into inactivity until the experiment had been finished.

Summary. The two most striking things shown in the data of Table 5 are (1) that the patient exhibits a tendency to be dominantly latent throughout the entire series with more marked latency immediately preceding and during the illness, (2) that on day 8/27 the patient's constant error takes a marked turn for the worse, furnishing what may have been a prognostic sign of her imminent somnolence. (See Lindley, 3*a*). No other behavioral or physiological changes had been noticed on this day.

VI. Psychometric Measurements

In order to obtain some measure of the patient's ability for tasks of a more complex nature, her responses to the Stanford-Binet tests were compared under somnolent and waking conditions.

In her normal state she scored a mental age of 17 years 3 months, IQ between 107 (16) and 123 (14). Below this level she failed the code, the problem of the enclosed boxes, and the difference between abstract words. At the superior adult level she passed the tests of rote memory for digits, of repeating digits backward, and of ingenuity. Her vocabulary was almost at the superior adult level.

The tests were administered again, during a somnolent period, about three weeks later. At that time a mental age of 12 years 9 months was obtained, with an IQ between 80 (16) and 91 (14). Because of her blurred vision it was not possible to give all parts of the test. The reliability of the total score suffers for this reason, although the tests that could be given were weighted to compensate for this lack. The basal year was 8 as compared with 14 at the previous testing. She failed in giving the date, weight discrimination, and repeating digits backward at the 9-year level, free association at 10 years, and the essential similarities test at 12 years.

A comparison of test responses from the Stanford-Binet and various other tests leads to the conclusion that the more complex processes were less disturbed by the drug than were the simple ones. Although there was a marked difference between the scores for rote memory of meaningless or disconnected material (digits, non-related words), there was little decrement during somnolence on immediate memory for meaningful material (connected paragraph, Healy-Bronner Learning D). The digit span, both backward and forward, was at 7 years during somnolence and at 18 years during waking, while memory for ideational material showed a minus difference of only 2 years instead of 9. Vocabulary scores were 72 and 66 respectively for the first and second lists, a relatively minor difference. Free association, however, which is probably a less complex process than the definition of abstract words, showed a marked retardation under somnolence. The unusually large score of 96 words in 3 minutes during the control test indicated that there was no paucity of associated material. During somnolence this score dropped to 36 words; all objects in sight were named and there was no organization into classes.

Three performance tests, the Seguin and Ferguson formboards and the Whipple-Healy tapping test, were given under each of the two conditions, but the marked motor retardation made valueless any

attempt at objective scoring. In general, it appeared that the higher levels of the Ferguson boards showed less decrement in skill during somnolence than the lower ones, the Seguin, or the tapping test.

Summary. In general, it may be concluded that the more complex abstract processes were only slightly diminished in efficiency during the period of luminal somnolence. The thought processes were slowed but remained logical, and both generalizing ability and visual mental representation seemed little disturbed.

General Summary

An examination has been made of the efficiency of a series of psychophysiological processes during somnolence induced by large doses of luminal in a patient addicted to episodic usage of the drug. The most severe disturbance was found at the lower "levels" of central-nervous-system activity: patellar reflex excitability, conditioned finger retraction to sound, simple attention maintenance, reaction-time to sound, rhythmical response to patterned sound stimuli, rote memory for digits and unconnected words, and fluidity of free association. Much less disturbance appeared at higher levels: complex attention maintenance, immediate memory for meaningful material, and abstract mental processes such as definition of words.

References

1. Bass, M. J. Differentiation of the hypnotic trance from normal sleep. *J. Exper. Psychol.,* 1931, **14**, 382-399.
2. Heiser, F., & Cohen, L. H. Diurnal variations of skin temperature. *J. Indus. Hygiene,* 1933, **15**, 243-254.
3. Lee, M. A. M., & Kleitman, N. Studies in the physiology of sleep: II. Attempts to demonstrate functional changes in the nervous system during experimental insomnia. *Amer. J. Psychol.,* 1923-1924, **67**, 141-152.

3*a*. Lindley, S. B. Reproduction of sound patterns by voluntary movements. (For comparative data obtained from normal subjects.) Ph.D. thesis, Yale Univ., 1933.

4. Miles, W. R. Correlation of reaction and coördination speed with age in adults. *Amer. J. Psychol.,* 1931, **43**, 377-391.
5. Pavlov, I. P. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. (Trans. & ed. by G. V. Anrep.) London: Oxford Univ. Press, 1927. Pp. xvi+430.
6. Scott, H. D. Hypnosis and the conditioned reflex. *J. Gen. Psychol.,* 1930, **4**, 113-130.

7. SEARS, R. R. & COHEN, L. H. Hysterical anesthesia, analgesia and astereognosis. *Arch. Neur. & Psychiat.,* 1933, **29**, 260-271.
8. SHIPLEY, W. C. Related studies in human conditioned-reflex learning. Ph.D. thesis, Yale Univ., 1933.
9. TUTTLE, W. W. Effect of sleep upon the patellar tendon reflex. *Amer. J. Physiol.,* 1924, **68**, 345-348.
10. WILSON, S. A. K. Modern problems in neurology. London: Arnold, 1928; New York: Wood, 1929. Pp. 372.
11. WOLFLE, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. Gen. Psychol.,* 1932, **7**, 80-103.
12. WOODROW, H. The effect of rate of sequence upon the accuracy of synchronization. *J. Exper. Psychol.,* 1932, **15**, 357-379.

Institute of Human Relations
Yale University
New Haven, Connecticut

MESURES PSYCHO-PHYSIOLOGIQUES PENDANT LA SOMNOLENCE CAUSÉ PAR LA PHÉNOBARBITALE (LUMINALE)

(Résumé)

On a examiné l'efficacité d'une série de processus psycho-physiologiques pendant la somnolence causée par de grandes doses de phénobarbitale (luminale) chez un malade a donné à l'usage de la drogue de temps en temps. On a trouvé le dérangement le plus sévère aux niveaux inférieurs de l'activité du système nerveux central: l'excitabilité du réflexe patellaire, la rétraction conditionnelle des doigts pour les sons, le maintien de l'attention simple, le temps de réaction aux sons, la réponse rhythmique aux stimuli conditionnels de sons, la mémoir machinale pour les chiffres et les mots seuls, et la fluidité de l'association libre. Beaucoup moins de dérangement s'est montré aux niveaux supérieurs: le maintien de l'attention complexe, la mémoire immédiate pour les mots pleins de significance, et les processus mentaux abstraits tels que la définition des mots.

COHEN, SEARS, LINDLEY, SHIPLEY, ET SNEDDEN

PSYCHOPHYSIOLOGISCHE MESSUNGEN AN TÄTIGKEITEN, DIE IN EINEM ZUSTAND DER DURCH PHENOBARBITAL (LUMINAL) VERURSACHTEN SCHLÄFRIGKEIT AUSGEFÜHRT WURDEN

(Referat)

Die Verfasser untersuchten an einem Kranken, der an den zeitweiligen Gebrauch grosser Dosen dieser Medizin gewohnt war, die Tüchtigkeit bei der Ausführung einer Reihe von psychophysiologischen Tätigkeiten unter

der Einwirkung der durch grosse Dosen Phenobarbitals (Luminals) verursachten Schläfrigkeit. Die schwerste Störung zeigte sich auf den niedrigeren Niveaux der Tätigkeit des zentralen Nervensystems. Es wurden also stark beeinflusst: die Erregbarkeit des Kniereflexes, die Schnelligkeit der bedingten Zurückziehungsreaktion des Fingers auf Schall [conditioned finger retraction to sound], Beharrung der einfachen Aufmerksamkeit [simple attention maintenance], Reaktionszeit auf Schall [reaction-time to sound], rhythmische Reaktion auf gestaltete Schallreize, Gedächtnis für Zahlen und für nicht-verbundene Wörter, und Fluidität der frein Gedankenassoziation [fluidity of free association].

Die auf höheren Niveaux stattfindenden Vorgänge erwiesen sich als viel weniger gehemmt. Es waren also weniger gestört: Beharrung der komplexen Aufmerksamkeit, Gedächtnis für frisch bemeisterten sinnvollen Stoff [immediate memory for meaningful material], und abstrakte geistige Tätigkeiten wie zum Beispiel das Definieren von Wörtern.

COHEN, SEARS, LINDLEY, SHIPLEY, UND SNEDDEN

A STUDY OF IMAGINED COLOR BLENDS*

From the Department of Psychology of the University of California at Los Angeles

KATE GORDON

It is a classical doctrine of empirical psychology that our minds do not create the sensory elements with which they operate in imagination. Imagination, even in its most productive guise, it is said, is but the rearrangement of old elements into new forms. This is not, however, a self-evident fact. Students are sometimes skeptical about it. It should be convenient therefore to have experimental evidence on the subject. This paper describes an attempt to demonstrate to class groups a way of studying their mental processes when they were asked to imagine the blends of certain given colors.

As an introduction we cannot do better than to quote from the works of Locke and Hume the following passages. John Locke in *An Essay Concerning Human Understanding* (1690), writing of simple ideas, said:

> But it is not in the power of the most exalted wit, or enlarged understanding, by any quickness or variety of thought, to invent or frame one new simple idea in the mind, not taken in by the ways before mentioned [i.e., by sensation or reflection upon sensation] nor can any force of the understanding destroy those that are there. The dominion of man, in this little world of his own understanding, being muchwhat the same as it is in the great world of visible things; wherein his power, however managed by art and skill, reaches no farther than to compound and divide the materials that are made to his hand; but can do nothing towards the making the least particle of new matter, or destroying one atom of what is already in being. The same inability will every one find in himself, who shall go about to fashion in his understanding one simple idea, not received in by his senses from external objects, or by reflection from the operations of his own mind about them. I would have any one try to fancy any taste which had never affected his palate, or frame

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, June 13, 1933.

> the idea of a scent he had never smelt; and when he can do this, I will also conclude that a blind man hath ideas of colours, and a deaf man true distinct notions of sounds.

In order to appreciate the force of Locke's contention we ought to reflect on the possibility of its opposite. It can hardly be said that there is anything illogical or repugnant to our thought in the idea that it *is* possible for the mind to generate ideas of colors or of sounds prior to the actual experience of them in sense perception. We certainly inherit an organism which is susceptible to these impressions. It is thinkable that some agitation of the cerebral cortex should arouse a consciousness of sense qualities in anticipation of perception. It was once held that we have innate ideas. We are told that, in the field of biological thought, the conception of spontaneous generation was held, in some quarters, up to the time of Pasteur's researches. The idea that there might be a spontaneous generation of simple impressions in the mind is not therefore an unprecedented view. I do not think that we could actually prove that the child *in utero* does not have them.

With regard to the priority of sense impression to the mental image we now quote from Hume's *A Treatise of Human Nature* (1738-48):

> Every simple impression is attended with a correspondent idea, and every simple idea with a correspondent impression. From this constant conjunction of resembling perceptions I immediately conclude, that there is a great connexion betwixt our correspondent impressions and ideas, and that the existence of the one has a considerable influence upon that of the other. Such a constant conjunction, in such an infinite number of instances, can never arise from chance; but clearly proves a dependence of the impressions on the ideas, or of the ideas on the impressions. That I may know on which side this dependence lies, I consider the order of their *first appearance;* and find, by constant experience, that the simple impressions always take the precedence of their correspondent ideas, but never appear in the contrary order.

Hume, however, has ingeniously thought up a possible exception to his own principle, and he presents it in the following hypothetical case:

> Suppose a person to have enjoyed his sight for thirty years, and to have become perfectly well acquainted with colours of all kinds, excepting one particular shade of blue, for instance, which it never has been his fortune to meet with. Let all the different shades of that colour, except that single one, be placed before him, descending gradually from the deepest to the lightest; it is plain, that he will perceive a blank, where that shade is wanting, and will be sensible that there is a greater distance in that place, betwixt the contiguous colours, than in any other. Now I ask, whether it is possible for him, from his own imagination, to supply this deficiency, and raise up to himself the idea of that particular shade, though it had never been conveyed to him by his senses? I believe there are few but will be of opinion that he can; and this may serve as a proof, that the simple ideas are not always derived from the correspondent impressions; though the instance is so particular and singular, that it is scarce worth our observing, and does not merit that, for it alone, we should alter our general maxim.

Now it is just this singular instance which seems to me interesting and important. I wonder whether Hume's hypothetical observer could really, after all, raise up to himself in imagination the missing link in the series of blues. It may be that Hume need not have made even this exception to his principle.

This example from Hume has been cited recently by Professor John Laird in his *A Study of Realism*. Laird expresses the belief that the case never could be put to experimental test because we never could be assured that the observer in question had not seen the particular shade of blue. But I think that an experimental approach can be provided to this question which need not assume that the observer has never seen the colors.

The greatest difficulty in answering Hume's query, as it seems to me, is to be found in this question: How are we to know just what *is* in the mind of the person who says that he can imagine or cannot imagine a certain color? How are we to inspect a sample of the color which he may profess to have? This difficulty must be clearly recognized; for it meets us at every point where we try to study the subject of mental imagery. We cannot communicate about colors unless we have a common system of reference, and the best way for me to tell you what color I am thinking of is to point to a physical

sample of that color which we both see under the same conditions of illumination.

The Munsell Company publishes a standard series of colored papers which they designate as "one hundred hues having the same saturation and varying by small steps." These color samples include red, orange, yellow, green, blue, and violet. They make a suitable scale for the purpose of our test. I pasted samples of these hundred colors on a large chart and numbered them as shown in Figure 1.

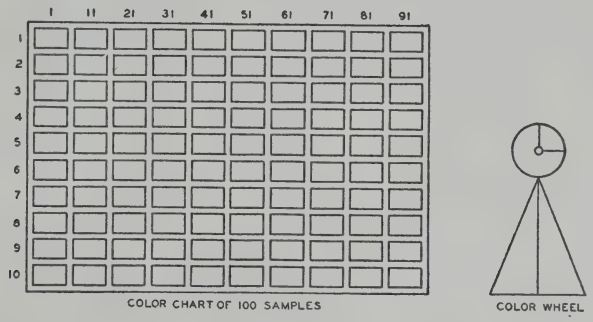


FIGURE 1

Each observer was provided with a mimeographed sheet containing the hundred numbers arranged as shown in Figure 2.

FIGURE 2

1	11	21	31	41	51	61	71	81	91
2	12	22	32	42	52	62	72	82	92
3	13	23	33	43	53	63	73	83	93
4	14	24	34	44	54	64	74	84	94
5	15	25	35	45	55	65	75	85	95
6	16	26	36	46	56	66	76	86	96
7	17	27	37	47	57	67	77	87	97
8	18	28	38	48	58	68	78	88	98
9	19	29	39	49	59	69	79	89	99
10	20	30	40	50	60	70	80	90	100

Any observer who held one of these mimeographed sheets could indicate a color on the chart by circling with his pencil the corresponding number on his mimeographed sheet.

I made up a series of color discs, also from the Munsell colored papers, for example, 180 degrees of blue-green with 180 degrees of

yellow. These were chosen by previous trial so that when rotated on the color wheel their blend should exactly match some one of the samples of color on the chart. I then set up the chart and the color wheel in front of a class of 103 persons, students of psychology.

The observers were asked to look at the first color disc and to try to imagine what the result should be when the disc was set in motion and the colors blended. They were asked to indicate their prediction by circling the number corresponding to the appropriate color on the chart. When these judgments or predictions had been recorded by the class, the color wheel was set in motion. During the rotation, while the actual color mixture was before them, the observers were asked to designate the number of the color which they then thought did match the mixture as perceived. By this means we could show the discrepancy, if there were any, between the color as predicted and the color as perceived. For instance, if the observer imagined that the colors on the wheel should look like sample 67 when they were mixed by rotation, and if, when he saw them actually mixed, he judged the blend to be like sample 55, then there would be a discrepancy of 12 steps.

The greatest possible discrepancy in this experiment is fifty steps. There are one hundred samples, but they form a color circle, the tones at the "hundred" end blending into those at the "one" end. Wherever the distance between two samples exceeds fifty, the reciprocal must be used. Thus, if a color is predicted to be like number 92, and is perceived upon rotation to be like 21, the difference is not 71 steps, but is 29 steps.

Typical results in the experiment group themselves into two opposite kinds, namely, those in which the predictions are fairly successful, and those in which they are quite blind. Where the colors to be mixed are analogous to one another, i.e., are close in the spectrum, the predictions are fairly good. Where the colors are far apart, i.e., approach the complementary relationship, the predictions are poor. In Table 1 two distributions of errors are shown. Column 1 gives the amount of discrepancy between a predicted and a perceived color blend. These amounts vary from zero to fifty, as explained above. Column 2 shows results for the blending of red 33 1-3%, yellow 33 1-3%, gray 33 1-3%. In column 2 the numbers stand for the numbers of persons (in a class of 103) whose predictions showed these discrepancies. Thus there were 16 persons

TABLE 1
IMAGINED COLOR BLENDS

Errors in prediction of color blends	Frequencies for red-yellow-gray	Frequencies for purple-yellow
0	16	3
1	18	0
2	23 M	2
3	16	2
4	9	2
5	4	1
6	3	2
7	3	1
8	2	1
9	2	1
10	2	3
11	0	0
12	0	2
13	0	2
14	1	2
15	1	0
16	0	0
17	0	0
18	0	0
19	0	0
20	0	4
21	0	1
22	0	1
23	0	2
24	0	4
25	0	3
26	0	2
27	0	2
28	0	1
29	2	5
30	0	6 M
31	0	5
32	0	4
33	0	3
34	0	2
35	0	3
36	0	3
37	0	0
38	0	1
39	0	3
40	1	5
41	0	3
42		1
43	*N* 103	1
44		0
45		5
46		2
47		1
48		1
49		1
50		2
		N 101

whose prediction was absolutely correct, i.e., the error was zero. There were 18 persons whose predictions were within one step of correct, and 23 persons who came within two steps of the correct color, this being the median group. In column 3 we have the results for the blending of purple 50% with yellow 50%, the numbers in the column standing for the frequencies of the errors in prediction. The predictions here are as wrong as possible within the conditions of the experiment.

It is probable in cases where the colors to be mixed are adjacent, as red and orange, that the observers, without necessarily having any image of such color, can predict that the blend will be an intermediate color. Several observers remarked as much. But where the colors are near complementaries it is not probable that the result can be reasoned out, and we see that in point of fact our observers were completely at loss to anticipate the result.

The feeling of surprise with which an observer often meets a color blend is a further evidence that these blends are not accurately anticipated or imagined. It is doubtful whether anybody, seeing a pair of complementary colors, could possibly anticipate the fact that they cancel each other. My experience is that observers sometimes fail to remember it even after they have watched the blending of complementaries on the color wheel.

In conclusion, I think we must acknowledge that so far as Hume's special instance is concerned we must still remain in doubt whether the interpolation of a color in a series is accomplished by means of a logical process or by a truly imagined color quality. We have, however, arranged a setting in which we may study objectively the ability to anticipate certain color blends. And we have found cases in which there is notable failure to anticipate correctly even though the elements of the blend are present to view and the resultant tone is also present for identification on the color chart. These cases tend to support the empirical position with respect to imagination.

University of California at
Los Angeles, California

UNE ÉTUDE DES MÉLANGES IMAGINÉS DE COULEURS

(Résumé)

On montre une méthode au moyen de laquelle on peut faire une comparaison entre (*a*) le résultat perceptible quand deux couleurs ou plus sont mélangées, et (*b*) le résultat prédit ou imaginé d'un tel mélange. Les divergences, s'il y en a, sont indiquées par degrés sur un tableau composé de cent échantillons de couleurs. Les résultats montrent qu'on peut faire une assez juste prédiction de la teinte de deux couleurs mélangées, quand ces deux éléments perceptibles sont adjacents dans le spectre. On peut supposer que ce résultat indique ou un processus de l'imagination ou un processus de la raison. Mais d'autres résultats montrent que les prédictions de la teinte sont très peu justes quand les éléments perceptibles sont bien loin l'un de l'autre dans le spectre, c'est-à-dire, quand ils s'approchent de la relation complémentaire. Cela nous permet de conclure que l'imagination, et la raison aussi, ne produisent pas pour nous une expérience nouvelle de couleur, quand même les parties constituantes seraient perceptiblement présentes et la teinte de ces parties constituantes serait aussi présente sur le tableau des échantillons. Il paraît à l'auteur que le résultat soutient la position empirique classique dans la psychologie, comme Locke et Hume l'ont exprimée.

GORDON

EINE UNTERSUCHUNG AN EINGEBILDETEN FARBENBLENDUNGEN

(Referat)

Es wird in Verfahren beschrieben welche eine Vergleichung ermöglicht zwischen (*a*) dem eigentlichen, perzeptuellen [perceptual] Ergebnis der Blendung zweier oder mehrerer Farben einerseits, und (*b*) dem vorausgesagten [predicted] oder eingebildeten Ergebnis einer solchen Blendung. Die etwaigen Abweichungen der beiden Ergebnisse von einander werden durch Abstufungen auf einer Farkenskala [chart] angedeutet, die aus 100 Farbenexemplaren besteht. Die Befunde erweisen, dass wenn zwei perzeptuelle Bestandteile [perceptual elements] im Spektrum nahe bei einander liegen, das Ergebnis ihrer Blendung ziemlich genau voraus gesagt werden kann. Dieser Befund deutet wohl entweder auf einen Einbildungsvorgang oder auf einen logischen Vorgang [process of reason] hin. Weitere Befunde, aber, weisen darauf hin, dass wenn die perzeptuellen Bestandteile bereiche darstellen, welche auf dem Spektrum weit von einander entfernt liegen, d.h., wenn ihr Verhältnis annähernd ein komplementäres ist [where they approach the complementary relationship], die Voraussagungen über das Ergebnis der Blendung ungenau sind. Durch diese Befunde wird der Schluss erlaubbar, dass weder die Einbildungskraft [imagination] noch die Vernunft [reason] im Stande ist, uns ein neues Farbenerlebnis [color experience] zu verschaffen, auch dann nicht, wenn die hierzu notwendigen Bestandteile perzeptiv anwesend sind und auch die Blendung dieser Bestandteile auf der Farbenskala zu finden ist. Der Meinung der Verfasserin nach, unterstützt dieses Ergebnis den von Locke und Hume erläuterten Gedankenkreis des klassischen Empirizismus in der Psychologie.

GORDON

SHORT ARTICLES AND NOTES

AN INVESTIGATION OF THE MECHANISM OF CALORIC NYSTAGMUS IN THE RABBIT

Joseph Gidansky[1]

Introduction and Historical Summary

The excitation of the semicircular canals by rotation of the head or by caloric stimulation produces, among other reflexes, nystagmus. This phenomenon appears in the eye in two phases, a quick phase and a slow phase, and the direction of the nystagmus is said to be that of the quick phase, although the slow phase represents a true compensatory movement. During the quick phase no visual sensation arises and therefore, when we have a nystagmus, space seems to turn around our body in the direction of the slow or opposite phase. The nystagmus, appearing in the case of rotation of the head, represents a rotation of the eye in the direction opposite to that of the head; therefore when we turn, as long as nystagmus is present, the space about us does not seem to move. The nystagmus is said to occur in the plane of rotation, that is, in the plane perpendicular to the axis of rotation; as we can turn the head in any position, the eye-nystagmus may be produced in any plane, but if the vertical axis is functioning the nystagmus will always be horizontal in space, although it will occur each time in a different plane of the head. Nystagmus is assumed to have three components, vertical, horizontal, and rotary, of which one or two may be lacking. In vertical nystagmus, the eye moves upward and downward; in horizontal nystagmus, the eye moves from side to side; while in rotary nystagmus, the eye rotates about the visual axis.

The nystagmus produced by means of stimulating the semicircular canals with cold water is different from that produced with hot water, and likewise the nystagmus produced when the head occupies different positions in space is also different. Bárány (1) was the first to notice both of these facts. He found that the direction of the nystagmus produced by the irrigation of one ear with cold water is directed towards the other side if the head is held in a backward position; but the nystagmus is directed towards the excited side of the head if the head is held in a forward position. Warm water produces the opposite results. His explanation of these facts is based

[1]The author wishes to acknowledge his indebtedness to Professor J. P. Nafe who made this study possible, and also to Dr. Raefel Lorente de No, Director of the Research Laboratory of the Central Institute for the Deaf, for his advice, assistance, and the use of his laboratory.

upon the assumption that as a result of local cooling of the labyrinth wall in one or more canals, in the case of cold-water nystagmus, endolymph currents appear, because the cooled endolymph will sink to the lowest part of the canal. This resulting endolymph current stimulates the end-organ of the ampulla. If the head of the animal is in such a position in space that the ampulla is higher than the cooled part of the canal, then the direction of the current is away from the ampulla; if the head of the animal is such that the ampulla is lower than the cooled part of the canal, the direction of the current is towards the ampulla. This same explanation holds for hot-water nystagmus, except that the direction of the endolymph currents is exactly opposite to that for cold-water nystagmus.

Buys (3) has carried out a very careful study of caloric nystagmus in humans and has recorded it by means of his nystagmograph. This study gives results which support the current-theory.

Bartels (2) criticized Bárány's theory. He was of the opinion that upon stimulation of the canals with cold water the labyrinth is shut off so that the nystagmus which appears is created in a way essentially similar so that spontaneous nystagmus appearing after extirpation of one labyrinth. The results of stimulation with warm water would be produced by a general raising of the activity of the excited labyrinth.

When Kobrak (7) found that it was possible to elicit the caloric reaction in a man with 5 cc. of water at 27° C., he believed that the effect of the heat upon the labyrinth was negligible. Thus Kobrak said that the cold or warm water produced a vasomotor effect. But three authors, Dohlman (5), Frenzel, and Schmaltz (13), at about the same time proved with very sensitive methods that even with so small an amount of water the heat or cold reached the labyrinth and produced a change in temperature in the semicircular canals, chiefly in the horizinal one in the neighborhood of the ampulla. This gave Bárány's theory a great deal of support. In a recent paper, Schmaltz (12) accepts the theory of Bárány as definitely proved.

DeKleijn (4) studied caloric nystagmus in rabbits and proved that the nystagmus in the rabbit, as in man, changes its direction when the head leaves the normal position and consequently allows the horizontal canal to deviate from its normal position. DeKleijn studied only one changing point of the nystagmus and concluded in favor of Bárány's theory. Lorente de No (8, 11) stimulated the individual canals and studied the reflexes of the isolated eye muscles. He found that the direction of the nystagmus in some of the muscles, chiefly the vertical and oblique ones, may be independent, under certain conditions, of the direction of the lymph current, because two different currents may produce the same nystagmus or one current may produce different nystagmus, depending upon

the position of the head in space, but when the lymph current changes its direction, the nystagmus, at least of a pair of muscles, always reverses. Lorente de No concluded that his experiments gave support to the theory of Bárány. Thornval (14), working with pigeons and studying the head-nystagmus, which displays the same phenomena as the nystagmus of the eye and eye-muscles, also concluded in favor of Bárány's theory.

Hartmann (6) tried to determine all the changing points in the various head positions in space of the nystagmus in the eye-muscles produced by caloric stimulation of the labyrinth. He found that the nystagmus of the horizontal muscles always changed its direction when the endolymph currents in the horizontal canal changed. With the vertical and oblique muscles, however, he found very complicated results, as sometimes the changing points were not separated by 180°, and, furthermore, they were different for the two pairs. It was not possible to correlate the changing points of these two pairs of muscles with corresponding changes of the endolymph.

Purpose of Present Investigation

The purpose of this study is to investigate the kind of nystagmus produced in the eye and in the six eye-muscles in various positions of the head in space with both monaural and binaural caloric stimulation of the vestibular apparatus. Particular emphasis will be placed upon showing the correlation between the inversion of the nystagmus and the endolymph currents in the semicircular canals. An attempt will be made to justify the conclusion that the vertical canals, in addition to the horizontal canal, at least the vertical anterior, also play a major rôle in the production of the nystagmus.

Method of Procedure

Rabbits were used throughout the experiment. They appear to be the most suitable animal because they display no reflexes to ordinary visual stimuli.

The apparatus used consists of a board about 2½ feet long and 1 foot wide, fitted with a Czermak clamp, so that the head of the animal may always be placed in the normal position with reference to the body and held in the same way throughout the experiment. The board, by means of an iron rod placed on its horizontal axis, was placed in an iron frame, which is supported several feet from the floor by a wooden stand. With this arrangement, all the different head positions may be obtained. The board, *B*, may be rotated around the axis *a-a'*, parallel to the bitemporal axis of the rabbit, and, similarly, by clamping the board tightly in place, the iron frame may be rotated around the axis *b-b'*, parallel to the longitudinal axis. The remaining head positions are obtained by adjusting

the board, for example, at different positions around the axis *a-a'* and rotating the iron frame. The degrees through which the rabbit is rotated are measured by means of a small double wheel-like instrument arranged with a needle, which freely rotates within the two wheels, which are fitted with spokes placed at 30° intervals. This device is screwed to the board, making possible the measurement of degrees of rotation about either axis.

The rabbit with its head in the clamp is placed upon the board in the normal position and bound tightly to assure no possible slipping during rotation. At the start, the head is always placed in the normal position, which makes the mouth form approximately a 35° angle with the horizontal plane and places the external semicircular canal horizontally in space.

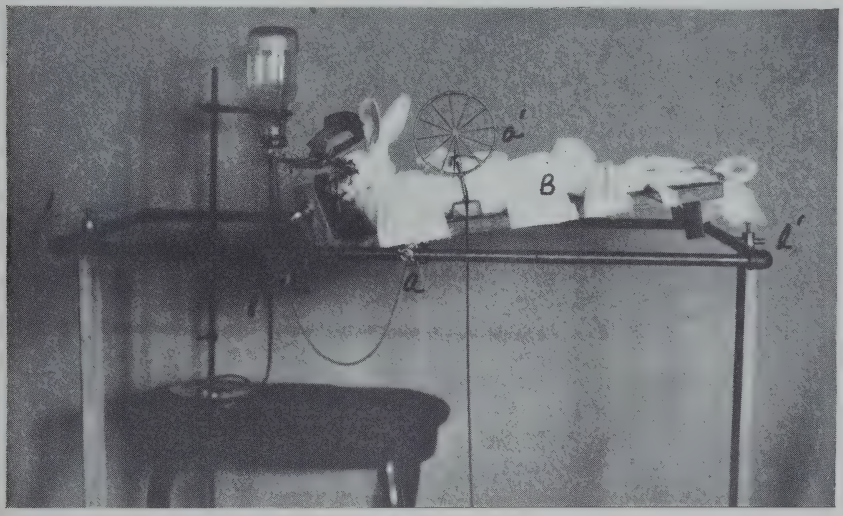


FIGURE 1
APPARATUS USED IN THE EXPERIMENT

The six eye-muscles are then prepared under cocaine anaesthesia, according to the Bartels technique (2). In this preparation of the eye-muscles, a thread is attached to the distal end of each muscle, which is then attached to the registering apparatus. This apparatus, designed by Lorente de No (9), as indicated by the diagram (Figure 2), consists of a crescent-shaped face, which is graduated so that four degrees are equal to a millimeter, and three levers, which make it possible to work with three muscles at a time. At the same point on the lever where the thread leading from the muscle is attached, rubber bands are also attached and are arranged to

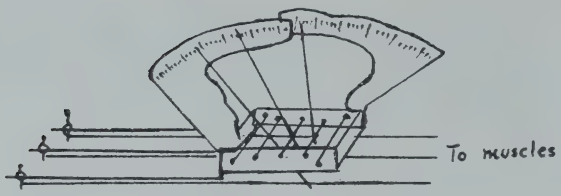


FIGURE 2
REGISTERING APPARATUS FOR MUSCLES

pull in the opposite direction. It is important that the muscles be stretched in their natural directions from their origins. When the muscles are at rest, at the beginning of the experiment, the rubber bands are adjusted so as to give a zero reading on the scale. When the muscle contracts or relaxes, it is always easy to read the amplitude of the muscle contractions upon the scale.

A copper tube about 3 inches long like that employed by Hartmann (6) was used in stimulation of the semicircular canals. An incision was made in the pinna, so that the copper tube could be introduced in the external canal as far as the drum membrane. This tube was constructed, as seen from the diagram (Figure 3), with one copper tube inserted into another. The water enters at *A*, circulates in the tube, and leaves at *B*. The tubes were fixed tightly in place against the temporal bone. Water approximately 10° C. was used for stimulation in order to bring about the reaction as

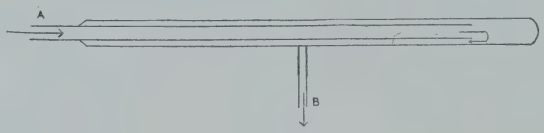


FIGURE 3
COPPER TUBE USED IN STIMULATION OF LABYRINTH

quickly as possible. The water was led from a flask, which was supported above the rabbit by a ringstand, by means of a rubber hose into the copper tube and drained through another rubber hose at the exit. As soon as nystagmus appeared the tube was withdrawn and replaced again against the bone in the next position. Stimulation was not continuous. This procedure proved to be a very efficient way to stimulate the labyrinth.

In several rabbits the changes in the eye-nystagmus were determined. The same method of procedure was used, except that, under cocaine anaesthesia, a cross was burned upon both eyes, by means of which the direction of the nystagmus was more easily obtainable.

Results

In the eye-nystagmus elicited by caloric stimulation, only one definite change in direction was found. This change occurs near the zero position, as can be seen from the table in Figure 5, in passing from the 30° position through zero to the 350° position. In going from the 30° position through the supine position (180°) to 330° position there is no sudden change in the direction of the nystagmus present. The diagrams in Figure 5 show that

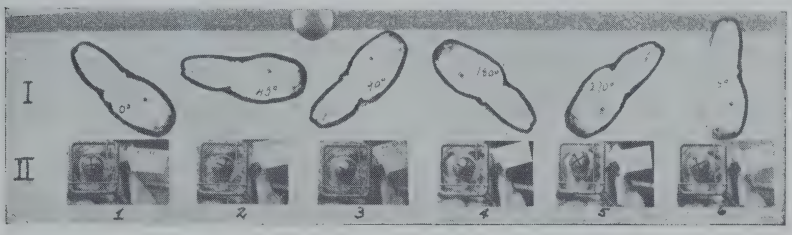


FIGURE 4

Positions of Head (I) and Eye (II) at Various Degrees

in passing from one position to the other the nystagmus slightly changes its plane, even in the neighborhood of the 180° position; during the whole turn, there is only a gradual change in the plane of the nystagmus. This is important to note because, since no nystagmus is present in the zero position, where the external canal lies horizontally in space, one would expect to find no nystagmus in the 180° position, where the horizontal canal again lies horizontally in space. As the eye in position 180° (Figure 4, *II, 4*) is

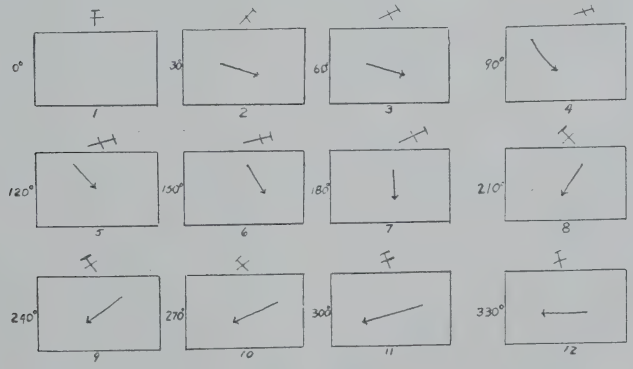


FIGURE 5

Direction of Nystagmus in the Eye at Different Head Positions
Cross denotes the position of the eye in the orbit; arrow denotes direction of nystagmus.

rolled forward in the orbit, it is very difficult to calculate which muscles take part in the production of the vertical nystagmus indicated in Figure 5, *7;* with certainty we can only postulate that the vertical and oblique muscles must show nystagmus.

The gradual change in the plane of the nystagmus, indicated in Figure 5, *2-12,* is not easy to explain. If we accept the fact that the caloric nystagmus is set up by the excitation of the horizontal canal, then we have the same lymph current from the 0° through 180° positions (Figure 5, *1-6*) as we have from 180° through 360° (Figure 5, *8-12*), except that in both cases the current is opposite. As the tonic reflexes determine the different eye positions for each head position (Figure 4), one is led to believe at first that the changes in the plane of the nystagmus are conditioned by the difference in the eye positions. This theory, however, would not be sufficient to explain all the facts; for example, in the 30° and 150° positions or in the 210° and 330° positions the eye has about the same position in the orbit, but the nystagmus is quite different.

Another possibility for explaining the changes in plane is to accept the fact that the nystagmus, as Lorente de No (11) demonstrated, is always produced by the combined activity of several muscles, and in each head position the amplitude of the nystagmus of the individual muscles changes. There does not seem to be any obstacle in the way of this assumption, which also agrees with the results of Hartmann (6). It is clearly seen in Figure 5, *6-8,* that the vertical and oblique muscles participate in the production of the nystagmus, and Lorente de No (11) has repeatedly stated that labyrinth stimulations resulting from the head position modify the amplitude and even the direction of the nystagmus of the muscles.

At any rate, the fact is that, although during the whole turn of the head the direction of the lymph currents in each of the semicircular canals must change two times, the eye-nystagmus shows a sudden change in direction only once, that is, in the neighborhood of the zero position, the normal head position.

From these results obtained in the study of the eye-nystagmus, no conclusive evidence was obtained to show the exact nature of the mechanism which produces the nystagmus, so we must proceed to the study of the eye-muscles themselves.

In the treatment of the data of the series of experiments with the eye-muscles, the division of the six eye-muscles in three pairs of antagonists will be used. The two horizontal muscles, that is, the internal and external recti, the two vertical muscles, the superior and inferior recti, and the two oblique muscles, the superior and inferior obliques, will be discussed separately. The nystagmus set up in the right (left) eye by stimulation of the left (right) labyrinth will be referred to in the discussion.

In the case of stimulation of the labyrinth on the same side the nystagmus has the opposite direction.

The Horizontal Muscles. Of the three groups of muscles, this group shows the strongest nystagmus with monaural stimulation with cold water in the many different positions in space about the bitemporal and longitudinal axes. At the zero position no definite nystagmus appears, but when the rabbit is turned only a few degrees downward a distinct nystagmus appears. There is a quick contraction of the external rectus with a simultaneous quick relaxation of the internal rectus. This same type of nystagmus continues approximately until the 180° position, where again no nystagmus is present. With the continued rotation, nystagmus reappears but in the opposite direction, the internal rectus now shows the quick contraction, and the external rectus the relaxation. This same nystagmus continues until the zero position is reached. The positions where there is no definite nystagmus are actually only a few degrees in the neighborhood of the 0° and 180° positions. The distance between the changing points of the direction of the nystagmus is approximately 180°.

The Vertical Muscles. This group differs greatly from the horizontal eye-muscles in that the nystagmus changes direction at approximately the 90° and 270° positions around both axes. The changes do not take place as sharply as in the previous group, actually occurring in a sector of about 15 to 20°. The superior rectus relaxes while the inferior rectus contracts quickly from the 0° to the 90° position, and the opposite takes place from there to the 270° position, where again a change occurs in the direction of the nystagmus. The changing points are approximately 180° apart.

The Oblique Muscles. The reaction in this group is considerably weaker than in the previous two groups; the contractions and relaxations are smaller and not as pronounced. The changes in the direction of the nystagmus take place at approximately the same positions at which the vertical muscles change. The distance between these two changing points is again approximately 180°.

It is important to note that the changes in the horizontal muscles occur quite clearly and quickly, while those in the vertical and oblique muscles do not occur as sharply, but in a sector of approximately 15° to 20°.

No nystagmus could be recognized when both labyrinths were excited simultaneously except when the rabbit was turned several degrees sideways. This is probably due to the lymph currents in the respective canals acting together, that is, reinforcing each other, instead of canceling each other's effects as in the former case where no nystagmus occurred. The work with binaural stimulation is extremely difficult and no conclusive results were obtainable.

Discussion

As was previously stated, the nystagmus of the horizontal muscles changes its direction suddenly when the horizontal canal passes through the horizontal position. This seems to prove that the direction of the lymph currents in the horizontal canal quite definitely determines the direction of this nystagmus.

A similar correspondence between the nystagmus of the vertical and oblique muscles and the lymph currents in the vertical canals seems also to be existent. Whether or not the change of temperature of the bony external ear is transmitted as far as the posterior canal cannot be determined without using special apparatus, but it seems to lie beyond any doubt that the vertical anterior canal must undergo the same changes as the horizontal one; both canals must be cooled off near the ampulla. According to this, the lymph current in the vertical anterior canal must change its direction in approximately the 90° and 270° positions.

The fact that the change of the direction of the nystagmus of the vertical and oblique muscles does not occur as suddenly as that of the horizontal muscles may be explained by two facts, both of which are certainly exerting their influence at the same time. In the first place, the portion of the vertical canal cooled off during the irrigation of the ear is undoubtedly greater than that of the horizontal canal, which gives a wider zone for inversion of the current. In the second place, we must consider the case when the vertical canal does not show any current because of the cooled point's being the highest or the lowest portion of the canal; the horizontal canal shows the strongest current because it lies vertically in space. This canal, as Lorente de No (10) has shown, produces nystagmus in all six of the muscles, which may be opposite in direction to the nystagmus set up by the excitation of the vertical canal. It is quite clear then that, as the head leaves the 90° or 270° position and the direction of the current in the vertical canal reverses, the direction of the nystagmus of the vertical and oblique muscles will not change at once, but will change just as soon as the excitation of the vertical canal overtakes the excitation of the horizontal one. The continuation of the same nystagmus of the horizontal muscles is explained by the fact that this canal has the strongest influence upon these muscles. This assumption seems to be a little complicated, but it is the only one that explains the results obtained in the case of nystagmus produced by rotation of the head.

The combination of the above-stated facts and the possible differences in the technique of applying the copper canal to the ear might explain the complicated results obtained by Hartmann.

Conclusions

1. The changes in the direction of the nystagmus in the eye and the inversion of the lymph currents in the semicircular canals are shown to be unrelated. During a complete turn of the head around the bitemporal axis there must be at least four changes in the direction of the lymph current, but the eye-nystagmus shows only one clear change in direction.

2. The nystagmus must be produced by excitation of the horizontal and of the vertical canals, at least of the vertical anterior.

3. The vertical and oblique muscles do not show as sudden a change in their nystagmus as the horizontal muscles because the change does not take place at the same moment that the lymph current in the vertical canal reverses its direction, but as soon as the excitation of the vertical canal overtakes that of the horizontal one.

4. Thus the muscles show as many changes in the direction of nystagmus as the lymph currents. The muscles also show nystagmus of very different amplitudes; however, the eye does not necessarily show the same phenomenon.

References

1. Bárány, R. Untersuchungen über den vom Vestibularapparat des Ohres reflektorisch ausgelösten rhymischen Nystagmus und seine Begleiterscheinungen. *Monatssch. f. Ohrenhk.,* 1906, **40,** 193-277.
2. Bartels, M. Über der Augenstellung durch den Ohrapparat. *Arch. f. Ophth.,* 1910, **76**, 1-.
3. Buys, E. Contribution à l'étude du nystagmus oculaire de la rotation chez l'homme. *Rev. d'oto-neuro-ocul.,* 1924, **2**, 641, 721; 1925, **3**, 10, 105.
4. de Kleijn, A., & Storm von Leeuwen, W. Über vestibulare Augenreflexe: III. *Arch. f. Ophth.,* 1922, 109.
5. Dohlman, G. Physikalische und physiologische Studien zum Theorie des kalorischen Nystagmus. *Acta oto-laryngol.,* 1925, **7** (Suppl.),
6. Hartmann, E. Experimentelle Untersuchungen über den kalorischen Nystagmus an dem isolierten Augenmuskeln des Kaninchens. *Acta oto-laryngol.,* 1927, **11**, 556-572.
7. Kobrak, F. Beiträge zum experimentellen Nystagmus. *Beitr. z. Anat., Physiol., Path., u. Therap. d. Ohres.,* 1918, **10**, 214-220.
8. Lorente de No, R. Die Grundlagen der Labyrinthphysiologie. *Skand. Arch. f. Physiol.,* 1926, **49**, 251-311.
9. ————. On the tonic labyrinth reflexes of the eye. *Acta oto-laryngol.,* 1926, **9**, 164-.
10. ————. Die Augewählte Kapitel aus der vergleichenden Physiologie der Labyrinthes. *Ergeb. d. Physiol.,* 1931, **32**, 73-242.
11. ————. The regulation of eye positions and movements induced by the labyrinth. *Laryngoscope,* 1932, **42**, 233-332.

12. Schmaltz, G. The physical phenomena occurring in the semicircular canals during rotatory and thermic stimulation. *Proc. Roy. Soc. Med., London,* 1932, **25**, 359-381.

13. Schmaltz, G., & Volger, G. Über die Temperaturbewegungen im Felsenbein bei der kalorischen Reizung des Vestibularapparates. *Pflüg. Arch. f. d. ges. Physiol.,* 1924, **104**, 708-717.

14. Thornval, A. Études experimentelles sur la fonction des organes des canaux semi-circular. Kopenhagen: Levin & Munksgaards Verlag, 1927. Pp. 103.

Washington University
St. Louis, Missouri

ON THE STRUCTURE OF ACTS

Donald K. Adams and T. L. McCulloch

Introductory

The behavior of animals may properly and profitably be considered to consist of acts. An act may be defined as all the behavior that is oriented toward a single goal of any order. Acts may have various degrees of unity, but they always have at least the unity of this reference to a single goal. Naturally, this does not exclude the existence of parts; and the unity of an act (beyond that implied by its reference to a single goal) will depend upon its internal structure, the mutual relations of its parts. The *identity* or the *character* of this structure may, under certain conditions, derive from that of its greatest part. Usually, of course, other factors than mere size (in terms of space, time, work, etc.) define its character.

In the present experiments the character of the act involved was primarily determined merely by the size, in the sense indicated, of its major part. Two T-shaped elevated tracks, 1 inch wide, were used. In each the "cross-bar" of the T was 66 inches (167.6 cm.) long. Thus in each the distance from the intersection to either end of the "cross-bar" was 33 inches (83.8 cm.). In the shorter track the "stem" of the T was 3 feet (91.4 cm.) long. In the larger track (which consisted of the shorter plus 24 feet of straight track) the "stem" of the T was 27 feet (822.9 cm.) long.

It would seem that the whole act of a rat in running the longer track might properly be characterized as "a long straight run ending in a right turning,"[1] and that the whole act of running the shorter track would

[1]Please note that this is different from "a long straight run *and* a turning." The behavior of the animals confirmed this characterization even before the end of the initial training period. The rats of the short-track group frequently oriented themselves toward the right as soon as placed on the track; i.e., their bodies were not parallel to the track, but the head was on the right and the tail on the left. The rats of the other group ran down the long track oriented straight ahead, in the direction of the intersection, until within 1 to 8 feet of that point.

approximate the character of merely "a right turning." If now two groups of rats be trained to turn to the right end of the "cross-bar," one group on each track, which group would have the greater difficulty in learning to turn instead to the *left* end? So far as we can see, no elementaristic theory could consistently make *any* prediction, though it would not be at all surprising if *ad hoc* hypotheses were offered *ex post facto* to explain either outcome. On the basis of any non-elementaristic theory, *Gestalt* or purposive, it could be (and was) predicted with confidence that the group on the longer track would have greater difficulty in shifting from a right to a left turning, since the change in the location of the food from the right to the left end of the "cross-bar" would transform the running of the short track into *a new act,* whereas the same change in the long track would make the running merely *the same act* (in which the right turning is embedded) with a difference. It is this proposition, i.e., that the change would constitute a new act in the case of the short track and not in the case of the longer, that would follow solely from non-elementaristic theories. The greater ease of shifting on the entirely new act would, of course, follow from any theory of retroactive inhibition; which phenomenon is here used merely as an indicator that, *for the rats,* the shift constituted a new act in the case of the short track and the same act with a difference in the case of the long track.

Two factors, to be sure, might militate against this issue (1) the very slight structuredness of the acts involved; and (2) the fact that the turning on the long track would be accentuated, as figure, upon the relatively homogeneous ground of the long straight run. If, in spite of these factors, the prediction should be fulfilled, it would seem to be *a fortiori* evidence for a theory of action wholes.

Experimental

The subjects of these experiments were 59 male albino rats obtained from the Albino Supply Company of Philadelphia. They were about two months old when purchased and were said all to have been born in the same calendar week. At the beginning of the experiments they were four months old. None had previously been used in any experimental work and physical living conditions were the same for all.

All of these animals were given ten days of preliminary training which consisted of running along *straight* elevated tracks of various lengths to food. At the end of this period all traversed the wooden strips with apparent facility. They were then divided into two groups, matched in respect to body weight.

Group I, containing 29 individuals, was trained to turn right on the shorter T and Group II, containing 30 individuals, to turn right upon the

longer T (which, it will be recalled, was simply the shorter plus 24 feet of straight track). Both groups were fed the same diet (calf-meal and milk) in amount sufficient to maintain them in good health. The reward consisted of a similar mixture, somewhat richer in evaporated milk than the regular diet and greatly preferred to the latter. Its composition was standardized throughout. A cup of this mixture was fastened to the right end of the "cross-bar" of the T and each animal was allowed to eat of it for 8 seconds upon arriving there. At the other end of the "cross-bar," the "wrong" end, a similar cup of the mixture was fixed; but this one was covered, invisibly except from directly above, by a piece of screen which barred the food from the animal but permitted the equilibration of olfactory stimuli from the two ends.

The animals were trained in batches of nine or ten. One rat was given a trial and then detained until each of the other members of his batch had received a trial before he was given another trial. The rats received from two to six trials per day, and were, of course, trained at approximately the same hour each day. On any particular day the members of both groups were given the same number of trials. Each rat was trained until he had obtained food on ten successive trials without error, i.e., without entering the left end of the crosspiece on which food was not available. He was then given 87 additional trials. Those rats which were tardy in meeting the criterion of ten successive runs without error were given a few additional trials each day, so that the training periods for all the individual rats were completed upon the same day. This training extended over a period of 37 days (April 21, 1933—May 27, 1933).

Test Period. On the day following the close of the training period the screen was shifted to the cup on the opposite side of the cross-bar. Food was now obtainable by the rat at the left end of the cross-bar and no longer obtainable at the right end. The rats were given two trials per day until they obtained food on the left on ten successive trials without entering the right half of the cross-bar.

Results. Table 1 presents a record of the number of errors made by both groups of rats in the training period and the number of trials given both groups, in terms of central tendency and dispersion. The averages of the number of errors made by both groups are almost identical. That is to say, both groups made approximately the same number of errors before consistently adopting the response of turning to the right for food. Table 1 shows that members of Group I were given, on the average, 7½ more trials than were members of Group II. This difference, though not large, may be taken to indicate that members of Group I were "overtrained" slightly more than were members of Group II.

Table 2 presents a summary of the performances of the two groups, in terms of trials and errors, during the test period. The table indicates

TABLE 1

SUMMARY OF RECORDS OF GROUPS I AND II DURING TRAINING PERIOD

	Errors		Trials	
	Group I (n=29)	Group II (n=30)	Group I	Group II
Mean	3.17	3.23	108.21	100.73
$\sigma_{dis.}$	1.98	2.98	2.26	3.77
Range	0-9	0-11	97-122	97-110

TABLE 2

SUMMARY OF RECORDS OF GROUPS I AND II DURING TEST PERIOD

	Errors		Trials	
	Group I (n=29)	Group II (n=30)	Group I	Group II
Mean	2.66	5.20	4.66	9.83
$\sigma_{dis.}$	1.37	3.55	3.12	6.32
Range	1-6	1-16	1-12	1-24
σ_{mean}	.2547	.6419	.5796	1.1452
Diff./$\sigma_{diff.}$, av.	3.74		4.03	

that the mean scores for Group I were 2.6552 for errors and 4.6551 for trials; for Group II the mean scores were 5.2000 and 9.8333 for errors and trials, respectively. These scores reflect a significant difference in the facility with which the members of the two groups adopted a consistent performance of the act of approaching the cross-bar and turning to the left for food. For error scores, the critical ratio, Diff./$\sigma_{diff.}$, av., is 3.7401; for trial scores it is 4.0344. These ratios indicate that the probability that there is a true difference between the performances of the two groups is almost unity.

A comparison of Tables 1 and 2 reveals that in the test period Group I made fewer errors than in the training period, namely, 2.6552 and 3.1724, respectively. The critical ratio, Diff./$\sigma_{diff.}$, is, however, only 1.1300, and is not considered statistically significant. Group II, on the other hand, made more errors in the test period than in the training period, the scores being, respectively, 5.2000 and 3.2333. The critical ratio is 2.3357. Although this value is ordinarily considered statistically insignificant, the probability that there was a true difference in error scores in the two periods is .9902. This difference between the differences seems to substantiate our inference that a shift on the shorter track makes it a new act, while a shift on the longer track makes it "the same act with a difference."

The main point, of course, is that the predicted difference in ease and speed of shifting occurs, and this is especially noteworthy in view of the

very slight structuredness of the acts. In the only other experiments we know of that bear even remotely upon this problem, those of Schwarz (1) upon human beings, much more complex acts, yielding a rich variety of complex structures, were involved. Many of the phenomena described by him, e.g., hesitation, wavering, recurrence of lapses after a pause, etc., occurred.

In the experiments just described the rats were run in the order: Group I (shorter path), Group II (longer path). Lest it be thought that this order might have something to do with results, we append the results of a similar investigation made somewhat more than a year earlier with smaller groups of rats of an entirely independent stock in another room and in the reverse order. That is, in the earlier experiments the long-path group was run before the short-path group. Tables 3 and 4 present the results of the preliminary investigation.

TABLE 3

SUMMARY OF RECORDS OF GROUPS A AND B DURING TRAINING PERIOD

	Errors		Trials	
	Group A (n=9)	Group B (n=6)	Group A	Group B
Mean	5.22	6.50	121.11	116.00
Median	4.00	2.50	102.00	104.00
Range	0-9	0-27	74-181	74-174

TABLE 4

SUMMARY OF RECORDS OF GROUPS A AND B DURING TEST PERIOD

	Errors		Trials	
	Group A (n=9)	Group B (n=6)	Group A	Group B
Mean	6.44	16.60	11.44	24.80
Median	5.00	15.00	7.00	26.50
Range	4-12	3-32	5-39	4-50
Diff./$\sigma_{diff.}$, av.	2.25		.9878	

It is further of interest to note that in this experiment we used a third group on a T-shaped track, the stem of which was of intermediate length, namely, 10 feet, and that in this group the animals tended to fall into two distinct subgroups, one of which approximated the long-path group in speed and facility of shifting to the left, while the other approached the short-path group in these respects. This suggests that these subgroups differed in respect to the way in which they experienced the running of

the track. Apparently for one subgroup it was "a long run ending in a turning," and for the other merely "a turning." The number of animals used was too small to permit stating this with certainty.

Summary

A situation was devised to test the proposition that the character of an act is determined by the relations of its parts to each other and to the whole act, and that under certain conditions the relation of their mere sizes might be the critical one. Thirty male albino rats were trained and overtrained to obtain food by going to the right end of the cross-bar of a T-shaped elevated track, the stem of which was 27 feet long. Twenty-nine matched rats were equally trained and overtrained to obtain food by going to the right end of the (same) cross-bar of a T-shaped elevated track, the stem of which consisted of the last three feet of the stem of the longer track. Upon completion of the overtraining, the food was shifted to the opposite end of the cross-bar, which required the rats of both groups to turn *left* at the intersection. It was predicted that the rats on the long track would discard the right and acquire the left turning more slowly than the rats on the short track since, if the above proposition is correct, the change in the locus of the food would imply, for them, a *change in the same act* while for the short-track group it would imply *a new act.* The rats on the long track required twice as many trials and made twice as many errors as did those on the short track in meeting the same criterion of performance in the changed situation.

Reference

1. Schwarz, G. Über Rückfälligkeit bei Umgewöhnung: Parts I and II. Parts IV and XVI of *Untersuchungen zur Handlungs- und Affektpsychologie,* ed. by K. Lewin. *Psychol. Forsch.,* 1927, **9**, 1-85; 1933, **18**, 143-190.

Duke University
Durham, North Carolina

PERSEVERATIVE TENDENCY IN MOTOR ACTIVITY

K. H. Rogers[1]

I. Introduction

Perseverative tendency is here defined as the effect, other than fatigue, which an activity exerts on a similar activity which immediately follows it. Using motor activity of a restricted nature, the present experimenta-

[1]The writer wishes to acknowledge his indebtedness to Professor S. N. F. Chant, Department of Psychology, University of Toronto, for co-operation and discussion respecting this study.

tion is concerned with interference and the carry-over of a speed set as possible evidences of this phenomenon.

A number of earlier experiments have had the study of perseverative tendency as their primary interest. The Dutch school of Otto Gross expressed appreciation of this phenomenon in its concept "secondary function," concerning which Weirsma (1906) and Heymans and Brugmans (1913) carried out experimental investigations. The work coming from C. Spearman's laboratory at the University of London (4) has offered, on the basis of careful experimentation, the suggestion that "perseveration" constitutes a second general function underlying cognitive activity. In this respect, the investigations of Lankes (1914), Wynn Jones (1915), Bernstein (1924), Pinard (3) (1932), and others are too well known to require further comment.

The recent work of Cathcart and Dawson (1) would seem to be a further contribution to this subject. Discussing the persistence effect or carry-over, of different rates that they found in working with an ergograph, these writers assert,

> It is as if there were in the psychophysical organism a kind of inertia, or general persistence of attitude, or "set," corresponding to physical inertia. This "carry-over" is what we have ventured to call "persistence."

This quotation expresses well the aspect of the more general problem to which the experimental work reported below is primarily directed. The subjects were 75 members of a third-year Arts class in Experimental Psychology.

II. Experimentation

Experiment I. Ascertaining the Effect of Interference on the Reaction-Time to Auditory and Visual Stimuli. The apparatus consisted of a choice reaction-time kymograph set-up for two sensory modalities, visual and auditory, and an experimenter's signal lamp (to control the presentation of stimuli) invisible to the subject and operated by a second synchronized kymographic arrangement.

The method involved three series of ten stimuli. An order to "reverse" (to change from responding to one stimulus with the forefinger of one hand to responding with that of the other hand) came after the first stimulus of each of the second and third series. The giving of stimuli was strictly according to prearranged order and intervals, in this manner avoiding a regularity of presentation and assuring constancy for all subjects.

Experiment II. The Carry-Over of a Speed Set in Signature Writing. A writing-board was used for performance and a kymograph hook-up for recording. The writing-board consisted of a brass plate set into one

end of a board, above which a thin sheet of spring brass was superimposed in such a way that a light touch upon the upper one would immediately bring the two into contact. Electrical connections were made from the plates to the kymograph markers.

The experiment consisted of two parts and required the subjects to come on two occasions. In each part the subject first wrote his name ten times in his "ordinary everyday fashion" using the writing-board. Then a presumed speed-set factor was introduced: the name was written ten times in an unusual manner with respect to speed, slowly in the first part of the experiment, and at a quickened rate in the second part. For writing slowly a metronome set at 30 (one tick every two seconds) was put in action before the subject; writing was done at the rate of one letter every two ticks. For fast writing maximum speed and effort were requested. These were not written on the board. The third set of ten signatures was written in the ordinary everyday fashion using the writing-board, in this way being as nearly as possible a repetition of the first set. So that we have in each part of the experiment two presumably similar performances with an unusual performance of the same activity interpolated between them.

Writing one's signature was selected for two reasons: it is a normal procedure and it is a type of behavior most likely to be uninfluenced by practice effect. Furthermore, writing it ten times to each set would overcome any initial lag that might present itself and secure a fair average of the time of performance. Subjects were instructed regarding the use of the writing-board and cautioned respecting possible causes of error.

III. Results

The scoring in both experiments was limited entirely to the time element. No attempt was made to evaluate the quality of the response in terms of errors, etc.

The effect on normal performance of the introduction of an unusual condition, as carried out in these experiments, is portrayed in Table 1, and Figures 1, 2, and 3.

TABLE 1

The Effect of Interference

Experiment	*N*	Basis of measurement	Av.	Sigma	Reliability Raw	Reliability Corrected
I	75	Increase after reverse signal	4.1 sec.	1.62	.70	.83
II	68	Increase after "slow" exercise	24.1 "	3.7	.80	.90
		Decrease after "fast" exercise	9.8 "	1.2	.69	.81

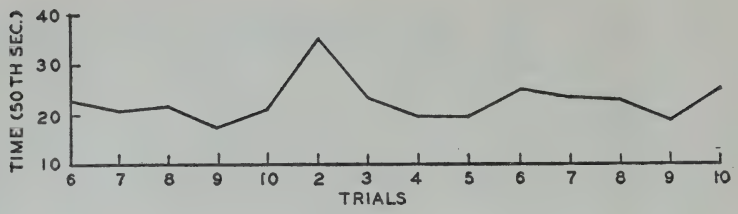


FIGURE 1

REACTION-TIME EXPERIMENT—THE AVERAGE REACTION-TIMES BEFORE AND AFTER REVERSE SIGNAL

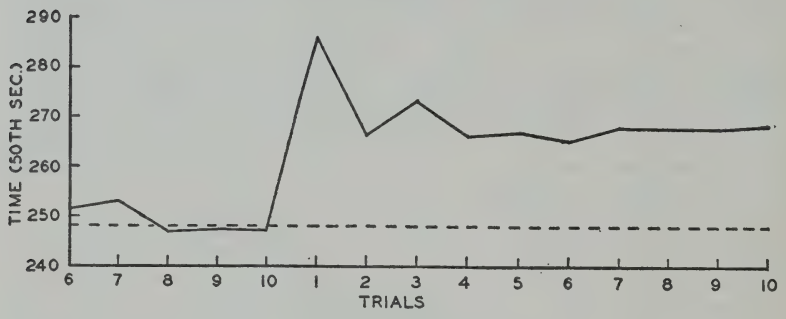


FIGURE 2

SIGNATURE-WRITING EXPERIMENT: PART 1—THE AVERAGE TIMES BEFORE AND AFTER THE "SLOW" INTERPOLATION

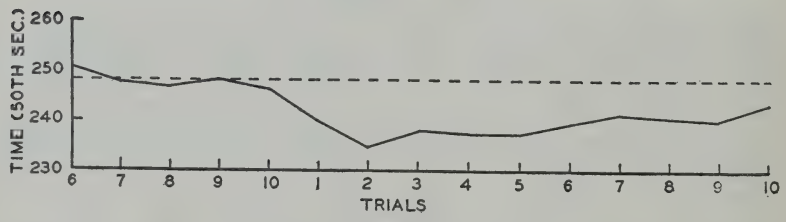


FIGURE 3

SIGNATURE-WRITING EXPERIMENT: PART 2—THE AVERAGE TIMES BEFORE AND AFTER THE "FAST" INTERPOLATION

In these experiments a significant degree of interference was plainly evidenced. Furthermore, the coefficients of reliability show the experiments to be reliable measures of disturbance; these were calculated on the split-half basis (odds *vs.* evens), and the corrections were secured with the Spearman-Brown formula for doubling the test.

The results were then intercorrelated. It was thought that this might possibly be suggestive respecting the claim that perseveration is of the nature of a general factor making for individual differences (4). The effect of "correction for attenuation" with these coefficients has been added.

TABLE 2

INTERCORRELATIONS BETWEEN EXPERIMENTS

Experimental items	r	r corrected	P.E. of r
Reaction-time and "slow" interpolation	.18	.24	±.087
Reaction-time and "fast" interpolation	—.19	—.27	±.084
"Fast" and "slow" interpolations	—.02	—.03	±.09

These coefficients (Table 2) mitigate against concluding the effect of a common underlying factor.

In order to compare the three types of responses with reference to the manner in which the "set" persists, curves were plotted. Figure 1 concerns the first experiment, and shows the data for the five responses before the reverse signal and the ten responses that followed the signal. Figures 2 and 3 show the curves for the two parts of the second experiment. Comparison indicates that, while Figures 2 and 3 show a marked resemblance, Figure 1 shows a curve that has little in common with them. This suggests a similarity between the "fast" and "slow" carry-over that is not present in the interference (reaction-time) situation; this, however, is to be contrasted with the coefficients of correlation.

IV. Discussion of Results

There is no question as to the significance of the carry-over effect as represented by the scores obtained. This is further supported by the evidence of the graphs. And the measures are shown to be reliable.

The lowness of the coefficient of correlation between the effect of the "fast" and of the "slow" activity is especially to be noted; considering the reliability of the results, this would seem to indicate that there is nothing in common between the two effects. This, however, must not be accepted without reservation, since the nature of the activity dealt with is such that the effect of the operation of a common factor may be cloaked. It was thought that possibly the fast writers—for it was found that the subjects varied greatly in speed of writing—would be affected more by the slow activity than by the fast, since they are nearer a fast limit; while the slow writers might show the reverse effect, since their writing, beginning nearer the speed of the slow activity, would show little effect here but would be subject to a considerable degree of "speeding-up." This, however, proved to be hardly the case. The fast writers are affected

little by either a slow or fast interpolation, while the slow writers are affected more by both, but particularly by the slow.

The intercorrelation results would suggest that there is no "common factor" or process operative in these three tests making for consistent individual differences. However, before definite assertions can be made in this respect, it would appear to be necessary to employ the essential criterion—the tetrad-difference criterion of Spearman, for example; this is hardly feasible with our limited results (4, pp. 72 f.).

It is just possible that these particular experimental settings for the general problem fail to isolate any "common" element. It is to be observed that "limit of response" activities are used throughout; in which case the "*g*" factor would be at a minimum and therefore any individual differences displayed would tend to be made up largely of the "specific" factor of each test situation. This suggests that there would of necessitiy be no significant correlation between the tests. That is to say, the measurement of "specific" factors has probably cloaked any "common" factor in the situation which necessarily must be very small. A further complication conditioning individual response is the effect of the element of "self-control." The experimental settings of the reaction-time experiment permits a greater influence of this, and its operation may well be detrimental to the display of a carry-over effect.

V. Introspective Accounts—Signature-Writing Experiment

Following each part of this experiment, the subject was asked to write a brief introspection. The chief interest of the experimenter had respect to whether the subject was aware of any difference in his "normal" writing following the set demanding an unusual mode of writing—slow or fast. There were 49 reports concerning part 1 (slow) and 47 for part 2 (fast). A brief summary of an analysis of these is given in Table 3.

According to these introspections it is to be observed that 38 cases, or 77.5%, and 41 cases, or 87.2%, are positively affected by some perseverative phenomenon.

This change might be either conscious or unconscious to the subject (by "conscious" is meant that the subject was aware of some change in his performance in the third set as compared with that of the first set of signatures). The results indicate that 26.5% were consciously affected in a positive direction in part 1 (slow) of the experiment, and 55.1% unconsciously; in part 2, 31.8% consciously, and 57.4% unconsciously, were so affected.

In these particular results it will be noticed that the effect is present though unconscious to a much greater extent than it is present and conscious. Also, the effect is present and unconscious in slightly more than 50% of the total number of cases in both instances. This would seem to

TABLE 3
ANALYSIS OF INTROSPECTIVE REPORTS

Part	N	Results*		Introspection†			Results compared with introspection							
							+		0		—		0	
		+	—	+	0	—	+	—	+	—	+	—	+	—
1	49	38	11	13	14	22	12	1	11	3	16	6	27	9
2	47	41	6	16	23	8	15	1	20	3	6	2	27	5

**Results*—the actual performance times recorded and measured.
+, those giving indicative results, i.e., indicative of perseverative tendency.
—, cases giving contrary results, i.e., the opposite effect from the preceding practice or set.
†*Introspection*—data culled from the introspective reports.
+, subjects who thought they were affected by the preceding set, i.e., when asked to write normally after writing slowly they unavoidably seemed to write slower than normally. And similarly, only conversely, with the fast writing.
0, those who were not aware of any deviation from the normal.
—, cases who thought they were affected negatively by the preceding set.

suggest that, as far as the phenomenon is shown in these experiments, it is predominantly of a physiological nature (distinguishing between mental and physical by the criterion of consciousness).

The physiological "automaticity" thus suggested is to some extent supported by the frequent occurrence in the introspective accounts of certain suggestive phrases: "a mechanical-like performance"; "a strong impulse to dash ahead"; "there seemed to be a resistance to moving quickly"; "it seemed unavoidable"; "it was difficult if not impossible to inhibit"; "it was against my will"; "my fingers insisted in hurrying after themselves"; "there was a tendency to race along"; etc.

These findings may be taken as indicative of perseverative tendency on the basis of agreement with one of Müller and Pilzecker's "proofs" (2) for such a phenomenon, namely, the third, which states that a continued motor activity sets up a tendency automatically to continue or to repeat the activity ("motorische Einstellung").

VI. CONCLUSIONS

1. The reaction-time experiment definitely indicates an interference effect; a possible manifestation of the phenomenon of perseverative tendency in motor activity. It is shown to be a reliable measure.

2. The signature-writing experiments definitely indicate the carry-over of a speed set; a possible manifestation of the phenomenon of perseverative tendency in motor activity. This test is shown to be a reliable measure

of the involved behavior. This conclusion appears to be affirmed by the analysis of introspective reports of the subjects.

3. It is not possible from these data to assert definitely either the presence or absence of a common factor or function underlying behavior in these experimental situations; that is, whether there has been measured a common perseverative tendency.

4. However, the low correlation between the three tests may be taken to suggest the absence of any functional unity. The indications of a similarity between the "fast" and "slow" carry-over that is not present between either or both of these and the interference phenomenon suggest that different phenomena are involved in the two experiments. It is felt that this might be precluded by the nature of the experimental situations.

VII. Summary

It was proposed to examine perseverative tendency as a characteristic of motor activity. The method was to determine the effect of an activity upon a similar activity immediately subsequent to it. Two experiments were carried out. The first, dealing with reaction-time, endeavored to relieve the phenomenon as an interference effect. The second experiment sought to appreciate this inertia process in terms of the carry-over of a speed set. Scores were secured in time units. The results indicate a positive effect throughout, and, while it is not possible to be conclusive respecting the relation between the phenomena set in relief by the experiments, the findings are sufficiently significant to warrant further research in this direction.

References

1. Cathcart, E. P., & Dawson, S. Persistence, a characteristic of remembering. *Brit. J. Psychol.,* 1928, **18**, 262-275.
2. Müller, G. E., & Pilzecker, A. Experimentelle Beiträge zur Lehre vom Gedachtniss. *Zsch. f. Psychol.,* 1900, Erg. 1. Pp. 300.
3. Pinard, J. W. Tests of perseveration. *Brit. J. Psychol.,* 1932, **23**, 5-19.
4. Spearman, C. Abilities of man: their nature and measurement. London & New York: Macmillan, 1927. Pp. xxiii+415.
5. Stephenson, W. Studies in experimental psychiatry: II. Some contact of p-factor with psychiatry. *J. Ment. Sci.,* 1932, **78**, 315-330.

University of Toronto
Toronto, Canada

AN EXPLANATION OF METFESSEL'S STROBOSCOPIC EFFECT OBSERVED IN MIRROR VIBRATION

NEIL WARREN

In a recent article Metfessel (1) has described a new type of stroboscopy in which the effect is produced in one of two ways: (1) the eyes of the observer are vibrated at an appropriate rate, or (2) a vibrating mirror reflects the image of the observed object. Either of these methods permits the intermittent vision of any given point on the retina necessary for stroboscopy. They vary, however, from previously described stroboscopic devices which involve one of the following methods: (1) the interruption of the light between the object and the eye of the observer, (2) the interruption of the light between the source of illumination and the object, (3) the interruption of the light at the object, and (4) the interruption of the illuminant. Since the introduction of the neon lamp the fourth method has been widely used. Use has been made of a vibrator, such as a tuning fork, in methods (1) and (2) for the purpose of interrupting vision by opening and closing a narrow slit. Certain facts indicate that the use of vibrations in the method described by Metfessel involves a principle different from these others producing the stroboscopic effect.

In the first place, vision as a whole is not intermittent in the sense of being completely interrupted part of the time. If a point of light is observed while the eyes are vibrated it is drawn out into a continuous line similar to a rapidly moving point, rather than appearing as one or more single points as would be the case if the eye were blinded part of the time. The same result is obtained when observing a point of light reflected by a vibrating mirror. The stroboscopic effect would seem to be due to the fact that at certain parts of the vibratory motion a relationship exists with the observed moving object which causes it to appear stationary. It is the purpose of this paper to determine at what point in the vibration the moving object seems to stand still, and to determine the conditions by which the most satisfactory result can be obtained. While it is probably safe to assume that the principles underlying both methods of vibratory stroboscopy are the same, this study deals only with the observation of a moving object in a vibrating mirror.

Three possibilities may be listed as explanations of the stroboscopic effect. It may appear stationary (1) during the movement of the mirror which is in the same direction as the movement of the observed object, or (2) when the movement is opposite to that of the object, or (3) at the points of change of phase where the mirror approaches a stationary position. If it could be assumed that the eye is completely blinded by the

movement this last would appear to be the logical explanation. That it is not, however, is indicated by the fact that the object is seen only once during a complete double vibration in which there are two changes of phase. A mirror vibrating at the rate of 50 d.v./second reveals the fifty lines on a stroboscopic disc which revolves once per second as single stationary lines. They cannot, then, be seen one hundred times per second as would be the case if seen during two momentary pauses at the change of phase. Moreover, the lines perpendicular to the direction of the vibrating mirror are not seen at all. Only lines moving in the same plane as the mirror may be stroboscopically observed. It is evident, therefore, that the moving object is seen either when the mirror moves in the same direction, in what might be called the positive relationship; or when the movement is in the opposite direction, the negative relationship.

In order to determine which of these two explanations is correct it is necessary to secure intermittent observation of the mirror so as to reveal the various phases of the vibration. This may be done by illuminating the mirror and moving object by a light flashing at the same frequency as the mirror vibrates. The same result is obtained if the observation is made through the openings of a revolving disc passing at the same frequency. Observation of the various phases of the mirror vibration in this way shows that the reflection of the moving object is stationary only when the direction of the mirror's movement is the same as that of the object. The lines of a revolving disc appear during the positive phase, but disappear during the negative phase.

The intermittent vision necessary for Metfessel's stroboscopy is obtained when the observed object is moving in a positive relationship to the mirror. It is probable also that the same effect is secured if the movement is in the same direction as the eye movement when the vibration is applied to the eye.

It is to be noted, however, that there is an optimal amplitude of vibration which produces the most definite stroboscopic effect. If the amplitude is too small the object appears faint and blurred. If the amplitude is too great, there is a spreading or doubling of the object. At the optimal point, the result is more satisfactory than that obtained by using a flashing light. A disc upon which are drawn radial lines will show all three stages of the effect for a given amplitude of vibration. Toward the center of the disc the lines will be double; halfway to the circumference they will appear as single black lines; nearer the circumference they will blur and fade out.

For the best result the amplitude of the vibrating mirror should be such that its greatest velocity is the same as the moving lines. In the example described above the speed of the lines at the circumference of the disc

is always greater than that of the mirror. Halfway toward the center of the disc the velocities are equal at the midpoint of the vibration. At the center of the disc the lines move more slowly, so that their speed equals that of the mirror both before and after the midpoint of the vibration.

The velocity of the moving lines is obtained by the following formula:

$$2\pi r/\text{sec.} \qquad [1]$$

in which r is equal to the radius of the disc.

The instantaneous velocity of the vibrating mirror is determined by the following formula:

$$2\pi fR \cos 2\pi ft \qquad [2]$$

in which f equals the frequency of vibration, R equals the amplitude, and t equals the time.

The lines are seen as stationary at the instant these two velocities become equal, that is when

$$2\pi r = 2\pi fR \cos 2\pi ft \qquad [3]$$

In order to secure the best results the amplitude of the mirror vibration should be such that the line is seen only at the midpoint of the vibration, so that

$$t = 0$$

in which case

$$\cos 2\pi ft = 1$$

and

$$2\pi r = 2\pi fR$$

Reference

1. Metfessel, M. Stroboscopy by means of impressed eye movements or mirror vibration. *Science,* 1933, **78**, 416-417.

University of Southern California
Los Angeles, California

INFLUENCE OF THE SPEED ATTITUDE ON TEST PERFORMANCE

Miles A. Tinker

With the origin and development of group testing in the fields of school achievement and intelligence it became necessary to choose between two methods of test administration. The available techniques were: (1) the work-limit procedure, or the time taken to complete a given task; and (2) the time-limit procedure, or the amount of work completed within a given number of minutes or seconds. Because of expediency, the time-limit procedure soon became the almost universal technique for administering group

tests. The time limits allowed for tests or parts of tests tended to be comparatively short (only 1½ minutes in some instances). This use of time limits has put an emphasis upon speed of response in most testing situations. The subject tends to adopt a speed attitude since he is made to feel that he is working against time. It is almost impossible, therefore, to give a standard achievement or intelligence test without setting up the speed attitude to some degree.

It is obvious that there are marked individual differences in rate of work. Is it possible then that rather strict time limits in testing are a handicap to some subjects and not to others? In an attempt to answer this question, test scores under standard time have been compared with scores under double and unlimited times. Some results (4, 5, 6) reveal little change in relative position of the subject in the group when the time limits are extended. In other experiments (1, 2, 3) the results demonstrate that certain subjects are considerably handicapped by the standard (restricted) time limits. Whether time limits should be extended, or the degree of extension desirable, has not been determined.

It has been assumed that, when unlimited time is given, the subject works in the absence of any speed attitude. When the time is merely doubled, however, it is probable that the speed attitude is still dominant since all test items will not have been completed to the satisfaction of most subjects. The present investigation is concerned with a comparison of test performance when the time limits are extended and the speed attitude is maintained with performance when the time limits are extended and a leisure attitude is dominant.

Two comparable groups of subjects were used. In one group were 110 freshmen and sophomores; in the other, 111. To the first group Form 5 of the Army Alpha (omitting Test 1) was given with the following time limits for Tests 2 through 8: 4, 1, 1, 1½, 2½, 2, and 3 minutes. These time limits, which are ½ to 1 minute less than the standard times, were employed to prevent more than a very few subjects from finishing before time was called. The subjects were instructed to work for speed as well as for accuracy. When the whole examination was completed, red pencils were distributed and the subjects were started again on Test 2. This time they were instructed to continue from one test to another until all the remaining items were done. At the end of the time equal to that previously used (single time), the subjects were told to leave that test and go on to the next, if they had not already done so. If they finished the last test ahead of time they were allowed to check their work until time was called. As before, they were instructed to work for speed and accuracy. In both the single (15 minutes) and double time (30 minutes), therefore, the speed attitude was dominant. No student was satisfied that

he had had sufficient time to complete entirely satisfactorily all items in the examination when time was called at the end of double time.

Form 5 of Alpha was given to the other group of subjects with the restricted time limits and the speed attitude as with the first group. When the examination was completed, the subjects were given red pencils and instructed that they would now have unlimited time to complete the rest of the items of the examination and to correct, if they wished, any of those already done. This method of setting up a leisure attitude appeared to be entirely successful. It was approximately two hours before the last test blank was handed in.

One week later the whole experiment was repeated with both groups and with Form 9 of Alpha. This provided a check on the first part of the experiment and also provided data to be used in computing reliabilities.

Scores and measures of variability are given in Table 1. It is apparent that the two groups are approximately equivalent in average scores and in variabilities. As is to be expected with a group of university students, the ranges of the scores are somewhat restricted. Increasing the time limits yielded a marked increase in average scores and a decrease in variabilities.

TABLE 1

SCORES AND VARIABILITY IN ALPHA TEST TAKEN UNDER SPEED AND UNDER LEISURE ATTITUDES

Method of administration	Form 5 Mean	Form 5 $\sigma_{dis.}$	Form 5 Range	Form 9 Mean	Form 9 $\sigma_{dis.}$	Form 9 Range
N=110			Speed attitude throughout			
Standard time	128.1	17.2	90	129.3	18.5	86
Double time	165.1	13.1	72	166.6	12.5	67
N=111			Leisure attitude for unlimited time			
Standard time	133.3	20.2	96	127.6	22.8	105
Unlimited time	173.2	14.1	64	171.6	14.2	64

It is interesting to note that there was almost as much increase in average score with double time as with unlimited time. The trends in the retest are the same as in the first test.

The reliabilities, computed by correlating scores on Form 5 with those on Form 9 given a week later, are all high. Increasing the time limit decreased the reliabilities slightly. The coefficients are listed below:

Reliability (N=110): speed attitude, standard time r=.863
Reliability (N=110): speed attitude, double time r=.819
Reliability (N=111): speed attitude, standard time r=.891
Reliability (N=111): leisure atittude, unlimited time r=.805

To discover the influence of the speed attitude in comparison with that of the leisure attitude under the extended time limits, the following tech-

nique was employed. Scores in single time were correlated with scores in extended time for each group. These coefficients were then corrected for attenuation. They are listed below:

Standard *vs.* double time, speed attitude, Form 5	r=.786
Ditto, corrected for attenuation	r=.929
Standard *vs.* double time, speed attitude, Form 9	r=.719
Ditto, corrected for attenuation	r=.849
Standard *vs.* unlimited time, leisure attitude, Form 5	r=.785
Ditto, corrected for attenuation	r=.927
Standard *vs.* unlimited time, leisure attitude, Form 9	r=.746
Ditto, corrected for attenuation	r=.881

The correlations between scores in standard time and in extended time are far from unity, even when corrected for attenuation. The homogeneity of the data tends, without doubt, to lower the correlations somewhat. Nevertheless, a few individuals apparently are handicapped by the restricted time limits. These results fall in line with those of Freeman (1, 2, 3). Trends are the same in the first and the retest data. In comparing the results under the speed attitude with those under the leisure attitude we find striking similarities. In both instances the extension of time limits yields about the same results. The correlations for scores on Form 5, between single and extended time, are .786 and .785 for speed attitude and leisure attitude respectively. For Form 9 (retest) the analogous correlations are .719 and .746.

Apparently it makes little difference whether a subject works with a speed attitude and with a somewhat restricted extension of time limits (double time, which is only 10 minutes longer than the standard time ordinarily employed for this examination) or with a leisure attitude and unlimited time. In general, a subject changes his level of performance approximately as much with a moderate extension of time and while working under pressure as with an unlimited extension of time which means working at one's leisure. In other words, the pressure of working with a speed attitude is no handicap to performance providing the time limits are not markedly restricted.

The implications of these findings are clear. For examinations of this type and with adult subjects, adequate results may be obtained when the subjects are working against time. The speed attitude as such is no handicap to performance. Time limits need be only long enough to permit all subjects to attempt most of the items. There is apparently no justification for allowing a subject to work on the examination as long as he wishes.

REFERENCES

1. FREEMAN, F. S. Power and speed: their influence upon intelligence test scores. *J. Appl. Psychol.*, 1928, **12**, 631-635.
2. ———. The factors of speed and power in tests of intelligence. *J. Exper. Psychol.*, 1931, **14**, 83-90.
3. ———. The factor of speed. *J. Gen. Psychol.*, 1932, **6**, 462-468.
4. MAY, M. Effect of doubling the time limits in the Alpha and Beta Examinations. *Mem. Nat. Acad. Sci.*, 1921, **15**, 415-420.
5. RUCH, G. M., & KOERTH, M. Power vs. "speed" in Army Alpha. *J. Educ. Psychol.*, 1923, **14**, 193-208.
6. RUCH, G. M. The speed factor in mental measurements. *J. Educ. Res.*, 1924, **9**, 39-45.

University of Minnesota
Minneapolis, Minnesota

WIDESPREAD IMPROVEMENT IN TACTILE DISCRIMINATION FOLLOWING LOCALIZED TRAINING

SHEPHERD IVORY FRANZ AND SYBIL KILDUFF

In carrying out a series of investigations on the relation of the cerebrum to tactile and visual perceptions a number of interesting facts were discovered, some of which have already been reported. Recently Mukherjee (5) has called attention to another item which has interest in the same field. He has reported a reduction in the double-point threshold accompanying localized training and a transfer of the effect of the training to other skin areas. A corresponding result for the localization of single tactile stimuli has been reported by Franz and Eaton (3) in the publications referred to above. Noteworthy transfer effects were also found when geometric forms were "learned" by skin areas (1, 2). Attempts have also been made, and are in progress, to analyze some of the conditions involved in tactile learning and transfer. One of these is briefly reported here, because of its relation to the problem of learning, and also because of its general confirmation of the results of Mukherjee and of the work reported in the other articles cited above.

Two linoleum blocks were cut into stimuli forms in the general manner described in one of the earlier articles (1). One of these was a straight line, 38 mm. long and 1 mm. wide; the other consisted of two corresponding lines of half length which met, forming an angle of 135 degrees. The stimuli forms were placed successively on the area of the skin being tested, and judgments were required. As a "normal" area the dorsal surface of the right forearm was used. Six young women subjects took part in the experiment. Because of a change in procedure with one subject, her results are not entirely comparable, and they are not reported

here although they show the same general effects as those from the other five subjects.

The subjects were informed that we were interested to discover how readily a person would learn to discriminate such forms as were to be applied to an area of the skin. They did not know that we were interested in the determination of "diffusion" effects. After some preliminary tests (from 80 to 160 for each subject) by which we determined the "normal" discrimination, training of the same area began. The training consisted in the repeated application of the two stimuli forms in succession, requiring judgments and correcting the judgments when necessary. After periods of training of this character (usually one week) tests were made on succeeding days to determine the results of the training. Immediately prior to each test series three applications of each of the two stimuli were made and the subject was informed of the sequences. The test series which was then carried on was conducted "without knowledge." In addition, after the completion of the normal and the training series, special tests were made (40 comparisons each) on three other body areas (contralateral forearm, anterior aspects of the homolateral and contralateral thighs).

The accompanying table shows the results on the normal (untrained) areas, those for the training series on the same areas, and the excentric transfer tests. Because there were only two stimuli used, chance would permit 50 per cent of correct judgments. Although this figure was not attained by all subjects in all preliminary series of tests, the general average is fairly close, and for WH the numbers of correct apprehensions were higher than 50 per cent in both of her series. The results for all subjects show that training improved the ability of discrimination to at least 90 per cent. Although the numbers of the excentric tests had to be kept small in order to prevent probable training of those areas, the results show on the average about 100-per-cent superiority over the "normal" untrained area, and they correspond with the final results of discrimination by the same area after it had been trained.

Excentric areas A, B, and C were not tested prior to the beginning of training of the right forearm. We did not wish to have these areas "trained" to any degree with the stimuli we used. It is a possibility that the final results for the three areas may be considered normal for them. On the other hand, this is quite unlikely in view of the gradual improvement in discrimination by the right forearm under training. Nor is it likely that the sensitivity of the left forearms (area A) for all subjects would be so much greater than the right as is indicated by the final results. Certainly, also, in other sensitivity tests the thighs have not been found to be superior in discrimination to the forearm. If the assumption is justified, as we think it is, that, had preliminary tests been

made on the excentric areas, results would have been obtained corresponding to those in the "normal" series, we are warranted in concluding that the training of one skin area has brought about a general effect in widespread segments of the body. If, moreover, we accept the conclusions of the clinical neurologists that the postcentral cerebral cortical areas (and adjoining regions) are concerned with tactile discrimination and localization, we are also forced to conclude that the effects of training are not localized segmentally, as clinical neurologists seem to believe, but that the effects are widely spread, both over adjoining and also over contralateral cerebral areas. This conclusion is in harmony with the conclusions from other results which have been referred to in the appended references. It is also consistent with conclusions which follow from studies of segmental visual learning (4).

TABLE 1

SHOWING IMPROVEMENT IN TACTILE DISCRIMINATION OF TWO FORMS ON RIGHT FOREARM WITH TRAINING, AND CORRESPONDING EXCENTRIC DISCRIMINATION WITHOUT TRAINING—*A*, ON LEFT FOREARM; *B*, ON RIGHT ANTERIOR THIGH; *C*, ON LEFT ANTERIOR THIGH

Forty tests were made in each series; the figures show the numbers of correct judgments.

Subjects	Normal				Training								Excentric		
													A	B	C
An	16	20			25	28	29	31	34	36	37	37	37	36	34
HI	17	16	17	15	19	22	25	31	34	36			37	35	34
HO	17	16	21		16	22	22	28	30	34	37		36	34	33
KN	16	16	17	22	17	21	35	26	32	35	38		33	33	29
WH	21	26			25	33	39	39					40	38	32

REFERENCES

1. FRANZ, S. I. Training in touch perception and cross education. *Publ. Univ. Calif. Los Angeles, Educ., Philos., & Psychol.*, 1933, **1**, 121-128.
2. ———. Diffusion effects following localized tactile training. *Publ. Univ. Calif. Los Angeles, Educ., Philos., & Psychol.*, 1933, **1**, 129-135.
3. FRANZ, S. I., & EATON, A. G. The possibility of training in tactile space perception. *Publ. Univ. Calif. Los Angeles, Educ., Philos., & Psychol.*, 1933, **1**, 117-120.
4. FRANZ, S. I., & MORGAN, R. G. Transfer of effects of learning to other retinal areas. *Publ. Univ. Calif. Los Angeles, Educ., Philos., & Psychol.*, 1933, **1**, 91-98.
5. MUKHERJEE, K. C. The duration of cutaneous sensation and the improvement of its sensible discrimination by practice. *J. Exper. Psychol.*, 1933, **16**, 339-342.

University of California at
Los Angeles, California

THE EFFECT OF COMBINING SOME GEOMETRICAL OPTICAL ILLUSIONS[1]

CHUNG-FANG NI

This paper reports an experiment in which three standard geometrical optical illusions were combined in pairs, and the effects of such combinations are compared with the separate effects of their two components.

The selected illusions are (1) the vertical-horizontal illusion; (2) the filled-unfilled space illusion; and (3) the feather-arrowhead illusion. A graphical representation of these illusions is unnecessary. The three combinations of 1-2, 1-3, and 2-3 are represented in Figures 1, 2, and 3 respectively. It is to be noted that each pair is so combined that the two illusions give the same direction of effect, i.e., they cause the same line to be overestimated.

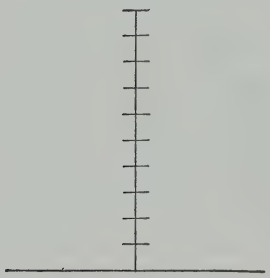


FIGURE 1
ILLUSION 1-2

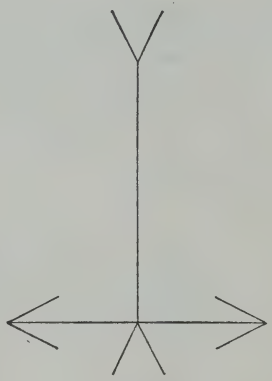


FIGURE 2
ILLUSION 1-3

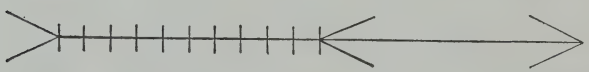


FIGURE 3
ILLUSION 2-3

These illusions have been extensively studied in isolation, but so far as the writer is aware they have never been studied in combination. Luckiesh (1, pp. 44-63) has referred to the combined influence of angles and

[1]From the Psychological Laboratory of the University of Chicago. The work was done under the direction of Professors Bills and Carr.

contrasting lengths, but he made no systematic study of their interaction. The combined effect of several illusions is often encountered in drawing and architecture. The problem thus has some practical as well as theoretical significance.

The method of right and wrong cases was utilized for measuring the degree of illusory effect. In each illusion the subject was asked to compare the relative length of two lines, one of which is underestimated and the other overestimated. Twenty-five drawings of each illusion were prepared, each on a separate card. These 25 drawings were all alike in respect to the underestimated line and all accessory lines. The length of the overestimated line varied from card to card by a difference of 1 millimeter. The set of 25 cards thus presented a gradated series in which the length of the variable line increased by 1-millimeter increments.

Each set of 25 cards was presented to each subject four times a day for five days, i.e., a total of 500 judgments was secured from each subject for each of the six illusions. In these 20 presentations of each set, the cards were so arranged that the variable line increased in length for the 10 odd-numbered trials and decreased in length for the 10 even-numbered trials. In order to minimize practice and fatigue effects the six illusions were presented according to the double-fatigue order. Four subjects were employed and a special exposure box was utilized so as to secure standard conditions of judgment.

The subject was asked to compare the length of the two lines and to give either a judgment of equality or inequality. The actual difference in the length of the two lines was ascertained for all judgments of equality and the average value of these differences was taken as a measure of the illusory effect.

The average value in millimeters and the P.E. for each subject and for each illusion are given in Table 1. The ratio of the combined illusions to the sum of the two components is given in percentage terms in Table 2. These comparative data justify the following statements:

TABLE 1

THE AVERAGE VISUAL ERRORS WITH THEIR P.E.'S FOR EACH OF THE FOUR OBSERVERS

The length of the standard line=100 mm.

Observers	Illusions					
	1	2	3	1-2	1-3	2-3
1	22.80±.27	7.89±.41	21.13±.31	28.23±.47	32.59±.36	17.80±.30
2	15.48±.59	6.02±.27	15.82±.47	18.29±.41	27.09±.19	11.57±.40
3	18.67±.35	5.16±.21	16.03±.40	21.71±.39	25.95±.27	10.91±.43
4	21.19±.59	6.83±.25	16.41±.61	24.06±.74	26.18±.67	12.70±.57

TABLE 2

PERCENTAGE RATIO OF COMBINED EFFECT TO ARITHMETICAL SUM

Observers	Illusions 1 and 2	1 and 3	2 and 3
1	92.0	74.2	61.3
2	85.1	86.5	53.0
3	91.1	74.8	51.5
4	85.9	69.6	54.6
Average	88.5	76.3	55.1

1. The effect of any pair of illusions in combination is always *less* than the sum of the two illusions separately. For subject 1, the combination of illusions 1-2 gives an average value of 28.23 mm., while the sum of the two values for 1 and 2 measured separately is 30.69. The combination effect of these two illusions is thus but 92 per cent of the sum of the two separate illusions—which is the value given in Table 2. We may thus conclude that any two illusions when combined do not give an *additive* or *summative* effect, i.e., the effect in combination is not equal to the sum of the two separate effects.

2. Two of the combinations, 1-2 and 1-3, give an effect that is *larger* than that of either component. This statement is true for all four subjects without exception. We may thus say that these two pairs exhibit a *facilitative* effect when combined. However, the degree of facilitating effect, as measured by the percentage values in Table 2, varies slightly from individual to individual, and varies markedly with the pair combined. For example, illusions 1 and 2 exert a more pronounced facilitating effect upon each other in combination than do figures 1 and 3, with the exception of subject 2 for whom the facilitating effect is practically equal for the two pairs.

3. The combination of illusions 2 and 3 gives a value that is *intermediate* between those of the two components. For example, the value for the combination 2-3 is larger than that for 2, and less than that for 3. This relation obtains for all four subjects. We may thus say that these two illusions exhibit a certain degree of *inhibitory* effect upon each other when combined and that the degree of this inhibitory effect varies slightly from subject to subject.

From the values for the three combinations one can *calculate* the effect of each component in the following manner. Let the calculated values of the three components be denoted by a, b, and c respectively. From the data of Table 1 we then have the following equations for subject 1: $a+b=28.23$, $a+c=32.59$, $b+c=17.80$. From these equations we can solve for a, b, and c, and these values for subject 1 are $a=21.51$, $b=6.72$, $c=11.08$. These

TABLE 3
CALCULATED VERSUS EMPIRICAL VALUES

Observers	Illusion 1			Illusion 2			Illusion 3		
	Empirical value	Calculated value	Percentage ratio	Empirical value	Calculated value	Percentage ratio	Empirical value	Calculated value	Percentage ratio
1	22.80	21.51	94.34	7.89	6.72	85.17	21.13	11.08	52.44
2	15.48	16.91	109.20	6.02	1.39	23.00	15.82	10.19	64.38
3	18.67	18.38	98.72	5.16	3.34	64.63	16.03	7.58	47.26
4	21.19	18.77	88.58	6.83	5.29	77.45	16.41	7.41	45.16
Average			97.71			62.56			52.31

calculated values for each component may be compared with the empirical values for the three separate illusions as given in Table 1. The comparative data for each of the subjects are given in Table 3, together with the percentage ratios of the calculated values to the empirical values.

With one exception, the calculated value is less than the corresponding value, i.e., the calculated effect of any illusion in combination is less than its separate effect. The calculated effect varies slightly with the individual, but it varies in a pronounced fashion with the illusion. Illusion 1 is most effective in combination, and illusion 3 the least effective. This calculated effectiveness in combination is not correlated with their effectiveness in isolation, for illusion 3 is nearly as effective as 1 in isolation but is the least effective in combination. Illusion 2 is much less effective than 3 in isolation but considerably more so in combination.

We may thus conclude that two illusions when combined do not sum their separate effects. When combined, they interact in an inhibitory manner, and the degree of inhibitory interaction varies with the pair combined. The calculated effectiveness in combination varies with the illusion, but this effectiveness is not correlated with their separate effectiveness. Any theory of the nature of these component illusions must account for the specific character of their interactive effects. Our data, however, contribute nothing definite to this end.

References

1. Luckiesh, M. Visual illusions. New York: Van Nostrand, 1922.

c/o Sing-Yen Book Store
Mah-Shih-Kew
Yunnanfu, Yunnan, China

APPARATUS

AN EXAMINATION OF ELECTRICAL-CURRENT-STABILIZING DEVICES FOR PSYCHOLOGICAL EXPERIMENTS

KARL F. MUENZINGER AND FRANK C. WALZ

The purpose of this paper is to communicate some experiences with the control of small currents in cases where large variations of resistance occur and to suggest some improvements in the method of control.

In certain experimental situations, e.g., the study of the galvanic skin reflex and the use of electric shock for punishment, it is essential that the current be as constant as possible. This requires that we should minimize the effects of variations of skin resistance upon the current. We may be concerned merely with the variations of any one individual subject or also with the variations between subjects. In the latter case the range of variations would naturally be greater. In either case it is necessary to know what range of variations to expect. To illustrate, with human subjects we may expect resistances as low as 15,000 ohms and as high as 65,000 ohms if we use mercury electrodes into which we dip a finger each. We would therefore have to take into account a variation of 50,000 ohms. With white rats (4) we may expect the minimum to be around 100,000 and the maximum well above 1,000,000 ohms if we use dry copper strips for the shocking grid. In this case we would have to take into account a possible variation of about 1,000,000 ohms.

A. The simplest method of minimizing the effects of variations in skin resistance is to use in series with the subject a fixed resistance of such a magnitude that the ratio between the variation in resistance to the fixed resistance is within our desired limit of accuracy. For instance, if we want to keep the current changes down to 5%, we would have to use an external resistance of 1,000,000 ohms for human subjects and 20,000,000 ohms for white rats for the set-up described in the preceding paragraph. The advantage of this type of stabilizing device is the simplicity of its design. Its disadvantage is the need for rather high voltages. If we were to use under the conditions described here a current of 2.0 milliamperes[1] for human beings and 0.2 milliampere for white rats the sources of our currents would have to have an E.M.F. of 2000 volts and 4000 volts respectively.

B. Instead of using a high external resistance, one may make use of

[1]All current notations in this paper are for direct current. These values must be properly transformed as pointed out elsewhere (5) in order to get comparable values for other types of currents such as alternating, full-wave, and half-wave rectified currents.

the fact, as first suggested for psychological apparatus by R. C. Davis (1), that under certain conditions the plate current of a vacuum tube remains fairly constant for large changes in plate voltage. If a variable resistance, such as the skin of a subject, is placed in such a plate circuit, its variations will produce changes in the voltage drop across the plate and only minor changes in the plate current. There are several methods of accomplishing this result:

1. In one of these methods the plate current is controlled by the filament voltage (2). However, radio tubes having either thoriated or oxide-coated filaments are not suitable for this purpose, because such filaments do not permit the use of the highest obtainable vacuum; thus, ionization currents are present, particularly at high voltages, making it impossible to obtain a true 'plate voltage saturation' and consequently a constant plate current with fluctuating plate voltages. On the other hand, a tube with a tungsten filament has a plate current in which the saturation level has a small slope, which means that a large change in plate potential produces a relatively small change in the current.

For our tests we chose the General Electric Radiotron UV202, which has a tungsten filament. With 300 volts applied to the plate and the filament voltage adjusted so as to produce a current of 2.0 milliamperes in the plate circuit containing a fixed resistance of 15,000 ohms (representing the minimum skin resistance of a person), the addition of 50,000 ohms (representing the maximum variation) caused a drop to 1.9 milliamperes, or a change of 5%. With 500 volts applied to the plate, a plate current of 0.2 milliampere, and a fixed resistance of 100,000 ohms in the plate circuit (representing the minimum skin resistance of a rat), the addition of 1,000,000 ohms (representing the maximum variation) caused a drop to 0.19 milliampere, or a change of 5%. We are able then to stabilize the current within the prescribed limits with voltage supplies of only 300 and 500 volts respectively, which otherwise would have had to be 2000 and 4000 volts respectively. In a motivation experiment (3) where the constancy of the current was essential, we used this tube with a plate voltage of 520 volts and a current of 0.15 milliampere. In this case a variation of 1,000,000 ohms in the skin resistance produced a current change of 3%, which is equivalent to using an external resistance of 33,000,000 ohms and a source of 5000 volts.

There are some serious disadvantages connected with this device to produce a constant current. In the first place, modern tubes having thoriated or oxide-coated filaments are unsuitable, as pointed out above, while the older tungsten-filament type receiving tubes are no longer on the market. In the second place, the current settles down to a definite value only after a certain period of adjustment or drifting. In order to overcome this difficulty in the motivation experiment just referred to we had to allow

5 minutes for the adjustment before using the apparatus and then we had to shunt the shocking grid periodically with a commutator so that the current flowed through the tube at least part of the time. In the fourth place, no rapid changes in current are possible on account of this factor of slow adjustment.

2. Another method consists in controlling the plate current through the grid voltage. We first tested a 3-electrode tube, 201A, which was contained in the circuit suggested by R. C. Davis.[2] With a plate potential of 250 volts and the current adjusted to 0.15 milliampere by means of a negative grid bias, the addition of 50,000 ohms to the plate circuit caused a drop of 40% in the current. A 3-electrode tube is clearly unsuited for our purposes.

We then tested some screen-grid tubes, 235, 57, and 58, which seem to have excellent characteristics. The current was adjusted in these tubes by varying the control grid bias. With 500 volts on the plate and 45 volts on the screen grid of tube 235, the addition of 1,000,000 ohms to the plate circuit caused a change of 7.5% in a plate current of 0.2 milliampere. In tube 57 the same situation produced a change of 7%. Tube 58 is similar to 57, but has the advantage of permitting the use of higher currents. The chief advantage of this device is that currents can be changed rapidly. It may perhaps be considered a disadvantage of both vacuum-tube devices that the plate current diminishes rapidly as the plate potential drops below 100 volts thus rendering these tubes unsuitable below this point. In order to overcome this difficulty the total E.M.F. in the plate circuit must be 100 volts larger than the voltage drop determined by the current used and the maximum resistance of the subject. For instance, if we want to use a current of 2.0 milliamperes with human subjects under conditions where the maximum skin resistance might be 100,000 ohms, the voltage drop due to these factors would be 200 volts; to this we would have to add 100 volts, so that the total E.M.F. for the plate circuit would have to be 300 volts.

3. The percentage changes in the last two devices are after all rather large for certain experimental requirements. In order to diminish these variations still further we suggest the following improvements for the use of screen-grid tubes. Since a decrease of a negative bias on the control grid causes an increase in the plate current, a grid potential which is taken off from a resistor in the plate circuit would in some measure com-

[2]We ought to call attention to the fact that Davis erroneously refers to this device as one that merely supplies a high resistance in convenient form (1). Its current changes are not a function of the resistance represented by the plate circuit but of the slope of the plate current. This correction is in no way intended to detract from the general value of Davis' suggestion.

pensate for a drop in the plate current. In order to increase this compensating effect we may even put an opposing E.M.F. into the grid circuit which would necessitate increasing the resistor in the plate circuit. In Figure 1 such a circuit is indicated for tube 57: *R* is the variable resistor

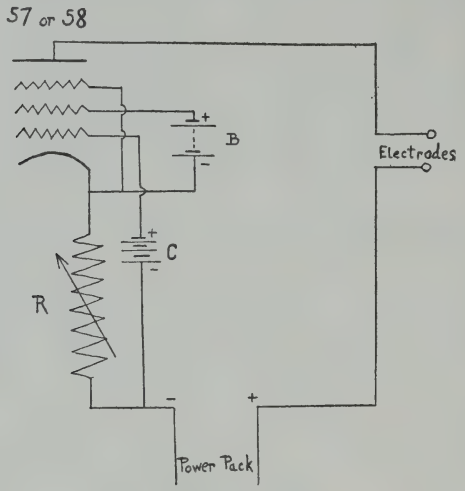


FIGURE 1

by which the plate current is adjusted; *C* is the opposing battery in the control grid circuit; *B* is the screen grid E.M.F. In Table 1 percentage changes of the plate current are given for cases where the control grid bias is supplied (*a*) from a battery, (*b*) from a resistor in the plate circuit, and (*c*) from such a resistor augmented through an opposing potential of 6 volts. The reduction in the variability of the current is quite striking under conditions (*b*) and (*c*). An entry of 0 in the table means that no movement could be observed in the needle of a microammeter having a full-scale deflection of 200 microamperes.

In conclusion, a few suggestions may be made as to the choice of the proper device for certain experimental situations on the basis of this table. In work with the galvanic skin reflex where it is of importance that a constant current pass through the electrodes, it would be best to use tube 57 with a plate E.M.F. of 300 volts, a screen grid E.M.F. of 90 volts, and condition (*c*) for the control grid circuit. This would provide an almost ideally constant current.

Where electric shock is to be administered to human subjects tube 58 might be used with a plate E.M.F. of 500 volts for a current up to 4.0

TABLE 1

Tube no.	Plate current milli-amp.	Resist-ance change ohms	Percentage change in plate current Screen grid 45 volts Condition (a)* %	(b)† %	(c)‡ %	Screen grid 90 volts Condition (a) %	(b) %	(c) %	Plate E.M.F. volts
235	.15	1,000,000	5.0	1.3	1.0		0.8		500
	.20		7.5	2.0	1.0		1.5	0.5	
	.05	100,000	0	0	0	0	0	0	300
	.5		2.0	0.5	0.3	1.8	0.3	0.2	
	1.0		5.5	1.5	0.8	5.0	1.2	1.0	
	2.0		14.0	4.0	1.0	30.0	4.0	4.0	
57	.15	1,000,000	6.0	1.2	0.1	3.3	0.15	0.05	500
	.20		7.0	1.8	0.3	5.0	0.2	0.1	
	.05	100,000	0	0	0	0	0	0	300
	.5		1.0	0.2	0	0.6	0	0	
	1.0		1.8	1.0	0.2	1.2	0.2	0.1	
	2.0		3.0	2.5	0.7	2.0	0.5	0.2	
58	.15	1,000,000	5.0	1.0	0.6	2.4	0.3	0.3	500
	.20		6.5	1.2	1.0	3.0	0.6	0.3	
	.05	100,000	0	0	0	0	0	0	500
	.5		3.0	1.2	0.4	0.8	0.1	0	
	1.0		6.0	3.0	1.0	1.0	0.2	0.1	
	2.0		7.0	3.5	1.0	1.0	0.5	0.2	
	3.0				2.5	2.7	1.0	0.3	
	4.0				3.0	4.0	2.0	0.8	

*Condition (*a*): control grid bias supplied from "C" battery.
†Condition (*b*): control grid bias supplied from resistor in plate circuit.
‡Condition (*c*): control grid bias supplied from resistor in plate circuit plus 6 volts opposing E.M.F.

milliamperes. For each additional 1.0 milliampere the plate E.M.F. should be increased by 100 volts. A screen grid E.M.F. of 45 volts combined with condition (*c*) seems to be sufficient to limit the current variations to 1%.

Where electric shock is to be administered to white rats tube 57 would seem best to use. Normally, shocks above 0.2 milliampere are not applied, so that a plate E.M.F. of 500 volts, a screen grid E.M.F. of 45 volts, and condition (*b*) would give a current variation of less than 1%.

References

1. Davis, R. C. A vacuum tube for stabilizing the current during measurements of the galvanic reflex. *Amer. J. Psychol.*, 1929, **41**, 474-475.
2. Davis, R. C., & Porter, J. M. A measuring device for the galvanic reflex. *J. Gen. Psychol.*, 1931, **5**, 115-120.

3. Muenzinger, K. F. Motivation in learning: I. Electric shock for correct response in the visual discrimination habit. *J. Comp. Psychol.*, 1934, **17**, 267-277.
4. Muenzinger, K. F., & Mize, R. H. The sensitivity of the white rat to electric shock: threshold and skin resistance. *J. Comp. Psychol.*, 1933, **15**, 139-148.
5. Muenzinger, K. F., & Walz, F. C. An analysis of the electrical stimulus producing a shock. *J. Comp. Psychol.*, 1932, **13**, 157-171.

University of Colorado
Boulder, Colorado

A UNIVERSAL FLEXIBLE ELEVATED FINGER MAZE

Ralph G. Whisler

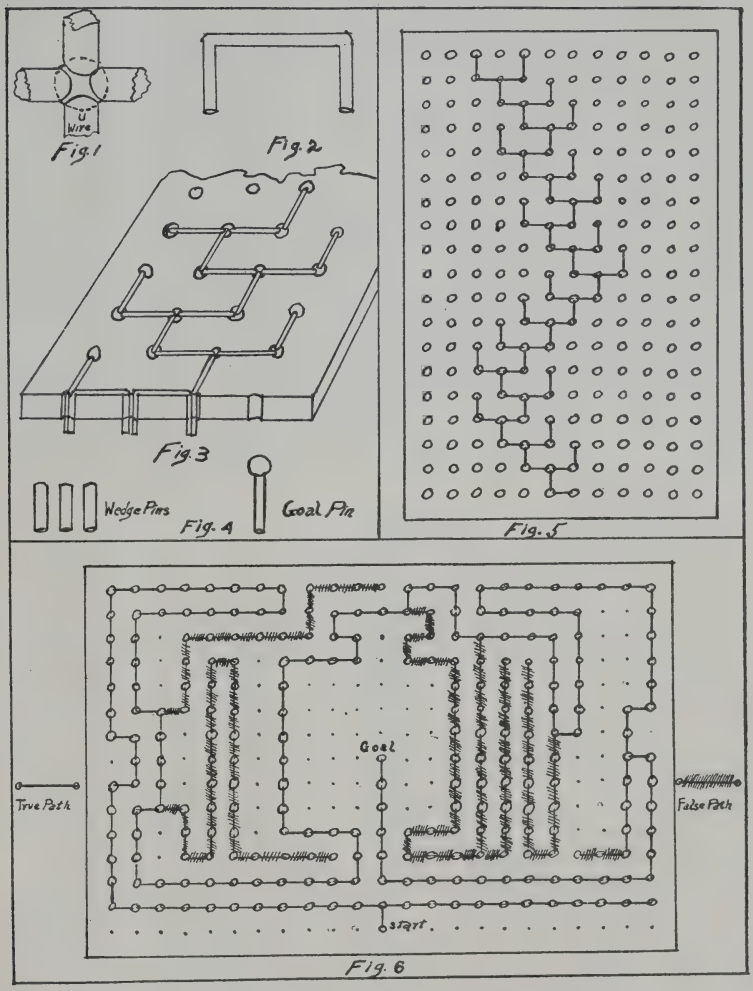
Recently the writer had need for a finger maze that was capable of being used in a number of situations closely following each other and where new maze patterns were desirable. To meet this need the maze herein described was devised, since the conventional finger maze is fixed permanently in pattern during its construction.

The initial impetus for the building of this apparatus was the desire to use a number of high-relief mazes of the Miles type for studies carried on by students in experimental psychology. The device, when constructed, enabled the experimenter to isolate and study a number of factors such as handedness, practice effects, transfer, and the like, as related to learning. The Miles maze, it will be recalled, is composed of a number of U-sections fastened into a base and with some joints soldered. The writer's maze apparatus also uses the U-unit, but with the difference that the units are not permanently fixed to the base as in the Miles high-relief finger maze. This renders the maze apparatus flexible enough to make possible the building-up of a great variety of the U- or multiple T-type mazes and also make possible the setting-up in the raised fashion of practically any of the patterns of the groove-, alley-, or slot-type mazes. This feature of flexibility also has the advantage that a new pattern can be set up in a few minutes' time thus making possible comparative studies in learning.

In setting up any maze pattern the paths and alleyways, whether true or false, are all represented in wire and a cul-de-sac is simply a wire coming to an end. At the goal a suitable marker is placed to indicate to the subject when he has arrived at his destination. As in all finger mazes, the subject works with the forefinger of his preferential hand, or otherwise at the experimenter's instruction, and is blindfolded, or the maze shielded from his vision.

The materials involved in the building of this device are relatively inexpensive, and this fact, coupled with the extreme flexibility of the apparatus, may recommend its use in a number of laboratories. The actual

materials used in the writer's model were a piece 3-ply ¼-inch panel stock and a quantity of number 10 galvanized fence wire. By the use of pliers, vise, hammer, forms, brace, and drill, the materials were worked into a very flexible device for the experimental study of learning or performance. The builder may choose his own material for the base and instead of using panel stock, which has the disadvantage of warping, may



substitute fiber, formica, bakelite, compressed masonite, or aluminum. The dimensions likewise may be varied at the discretion of the builder.

The writer's apparatus was constructed as follows: A piece of panel stock, ¼ inch thick, and 24 inches by 36 inches in size, was laid off and drilled at the corners of 1½-inch squares with a ¼-inch drill. If a different sized wire is used the holes should be made large enough to accommodate four of the wires quite snugly (Figure 1). The base when drilled contains 345 holes the center of each one being 1½ inches distant from its nearest neighbor. The wire is cut in lengths of 2½ inches and so bent as to form U-shaped pieces (Figure 2). The inside measurement of the wire is made equal to the distance between the proximal sides of adjacent holes so as to fit tightly (Figure 3). In the model which was constructed, the inside length of the U was 1¼ inches, since this was the distance between the proximal edges of the holes connected by the U-wire. It is desirable in the interest of facilitating the building and the changing of patterns to make the length of the U-legs about ⅜ inch longer than the thickness of the stock used for the base. To take care of the extension of the U-legs below the base, it is necessary to build the base on small rubber stoppers about ½ inch high.

The wire may be cut with a pair of wire cutters or a chisel and bent in a vise or over a form built for the purpose. If desired, a variety of U-lengths may be made, keeping in mind always that the inside length is governed by the distance between the proximal edges of the holes to be connected. To this, of course, must be added enough length to provide for the U-legs.

It is quite essential that all laying-out, drilling, bending, and cutting be done very accurately in order for all parts to be interchangeable. If larger wire is employed the holes must be enlarged and the U-lengths shortened accordingly, the reverse being true if smaller wire is used. It will be found in constructing various maze patterns that not all holes will have four paths leading out of them, but this will offer no obstacle to the use of the apparatus if it has been properly constructed. It is desirable, however, to have at hand a supply of short wire lengths, about ½ inch long, to fill in the holes and to take up play due to possible inaccuracy and wear and thus to make the pattern rigid. If wedge pins (Figure 4) are found necessary they should be pushed well below the surface of the maze pattern to avoid any possibility of the subject's receiving any tactual cues while tracing the maze.

The writer believes that this maze apparatus is sufficiently flexible to meet many conditions that are not met by the conventional fixed type of finger maze. For example, the standard type multiple-T maze may be constructed (Figure 5) in as many forms and arrangements as may be desired. In addition, the pattern of the standard fixed maze of the groove or pencil type may be translated into the elevated form. As an example of this (Figure 6) we may take the familiar Hampton Court maze.

The Ohio State University
Columbus, Ohio

AN ELECTRICAL MAZE STYLUS

RALPH W. GILBERT

In a recent experimental study of human maze learning it was desired to administer shocks when blind alleys were entered without the use of an extra electrode applied to some part of the subject's body other than the working hand. To accomplish this, the stylus described below was used. It has been in use for two years and has proved itself to be quite satisfactory.

The tip of the stylus is of brass, 3/16" in diameter, turned down from the brass base of the handle, which is about ½" in diameter. The upper part of the handle consists of a hollow bakelite core, around which are fitted about 30 rings of brass and bakelite, arranged alternately. These rings are about ⅛" wide. All connections are made internally. One of the three flexible wires is connected with the tip, which is insulated from the rings, while each of the other two wires is connected with alternate brass rings of the handle; i.e., half the number of rings is connected with one wire, the remainder with the other. When in use, the tip and its attached lead constitute a segment of the primary circuit of the inductorium, while the leads from the rings are connected with the secondary terminals.

FIGURE 1

The maze used has brass contact plates at the ends of the blind alleys, all connected by wires concealed in the base, with a common binding post. When the maze, inductorium, source of primary current, and stylus are connected in series, the primary circuit is closed whenever the stylus tip comes in contact with any one of the blind-alley plates. The handle is so designed that it cannot easily be grasped without bringing the hand in contact with at least two brass rings, thus making the administering of the shock a positive affair.

While this stylus was designed with the maze and the Porter inductorium as the source of shocking current in mind, it might easily be adapted for use in connection with other pieces of apparatus, such as special mirror-tracing set-ups, or with other sources of electrical stimulation.

New York University
New York City

BOOKS

ANASTASI, A. Practice and variability: a study in psychological method. *Psychol. Monog.*, 1934, **45**, No. 5 (Whole No. 204). Pp. 55.

BETTS, E. A., & BONTRAGER, O. R. Research studies in elementary school language, No. 1: An evaluation of certain techniques for the study of oral composition (Betts). An experimental appraisal of pupil control of certain puctuation items (Bontrager). *Univ. Iowa Stud.: Stud. Educ.*, 1934, **9**, No. 2 (New Ser., No. 271). Pp. 62. $.50.

CALIVER, A. Secondary education for negroes. *U. S. Dept. Interior, Off. Educ., Bull.*, 1932, No. 17. (*Nat. Survey Sec. Educ., Monog.* No. 7.) Washington, D. C.: U. S. Govt. Print. Off., 1933. Pp. ix+121. $.10.

COOK, T. W. Binocular and monocular relations in foveal dark adaptation. *Psychol. Monog.*, 1934, **45**, No. 3 (Whole No. 202). Pp. vi+86.

CORNELL, E. L., & COXE, W. W. A performance ability scale: examination manual; individual record blank. Yonkers, N. Y.: World Book Co., 1934. Pp. iv+88+4. Manual, $1.50; blanks, pkg. of 25, $.90.

ENGELHARDT, F., ZEIGEL, W. H., JR., & BILLETT, R. O. Administration and supervision. *U. S. Dept. Interior, Off. Educ., Bull.* 1932, No. 17. (*Nat. Survey Sec. Educ., Monog.* No. 11.) Washington, D. C.: U. S. Govt. Print. Off., 1933. Pp. viii+207. $.15.

EY, H. Hallucinations et délire: les formes hallucinatoires de l'automatisme verbal. Paris: Alcan, 1934. Pp. x+192. 15 fr.

FERRIÈRE, A. Der Primat des Geistes. Langensalza, Berlin, & Leipzig: Beltz (n. d.). Pp. viii+260. Rm. 7.50.

FREEMAN, G. LA V. Introduction to physiological psychology. New York: Ronald Press, 1934. Pp. xviii+579. $4.50.

GENGERELLI, J. A. Brain fields and the learning process. *Psychol. Monog.*, 1934, **45**, No. 4 (Whole No. 203). Pp. 115.

GUNN, M. A. A technique for improving basic skills in English in high school. *Univ. Iowa Stud.: Stud. Educ.*, 1934, **8**, No. 7 (New Ser., No. 266). Pp. 40. $.50.

HARTSHORNE, C. The philosophy and psychology of sensation. Chicago: Univ. Chicago Press, 1934. Pp. xiv+288. $3.00.

IOWA CHILD WELFARE RESEARCH STATION. Manual of nursery school practice. *Bull. State Univ. Iowa*, 1934, New Ser., No. 730. Pp. 215. $1.00.

KEESECKER, W. W., & SEWELL, F. C. Legal and regulatory provisions affecting secondary education. *U. S. Dept. Interior, Off. Educ., Bull.*, 1932, No. 17. (*Nat. Survey Sec. Educ., Monog.* No. 9.) Washington, D. C.: U. S. Govt. Print. Off., 1933. Pp. vi+114. $.10.

KEFAUVER, G. N., NOLL, V. H., & DRAKE, C. E. The secondary-school population. *U. S. Dept. Interior, Off. Educ., Bull.*, 1932, No. 17. (*Nat. Survey Sec. Educ., Monog.* No. 4.) Washington, D. C.: U. S. Govt. Print. Off., 1933. Pp. viii+58. $.10.

LAHKOVSKY, G. Le racisme et l'orchestre universel. Paris: Alcan, 1934. Pp. 153. 15 fr.

LOOMIS, A. D., LIDE, E. S., & JOHNSON, B. L. The program of studies. (In two parts.) *U. S. Dept. Interior, Off. Educ., Bull.,* 1932, No. 17. (*Nat. Survey Sec. Educ., Monog.,* No. 19.) Washington, D. C.: Govt. Print. Off., 1933. Pp. x+340. $.15.

MARKS, E. S., & HOROWITZ, E. L. (with the assistance of SEWARD, J. P., & MURPHY, G.) Psychology work-book, to accompany Gardner Murphy's *General psychology.* New York & London: Harper, 1933. Pp. 97. $.50.

PRADINES, M. Philosophie de la sensation: II. La sensibilité élémentaire (les sens primaires). Les sens de la défense. (*Publ. de la Faculté des Lettres de l'Univ. de Strasbourg,* No. 66.) Paris: Les Belles Lettres, 1934. Pp. 381. 50 fr.

PRESSEY, S. L. Chapter tests to accompany *Psychology and the new education* by S. L. Pressey. New York & London: Harper, 1934. Pp. 54. $.30.

PRESSEY, S. L. Questionnaires to accompany *Psychology and the new education* by S. L. Pressey. New York & London: Harper, 1934. Pp. 18. $.10.

TUTTLE, H. S. A social basis of education. New York: Crowell, 1934. Pp. x+589. $3.00.

VITELES, M. S. The science of work. New York: Norton, 1934. Pp. xii+442.

WOODWORTH, R. S. Psychology. (3rd ed.) New York: Holt, 1934. Pp. xiv+546. $2.50.

SUBJECT INDEX

Abnormal and anthropology, 59
Absolute tempo, 230
Acts, structure, 450
Age and intelligence scores, 208
  and the kinephantom, 204
Algebra of luminosity, 286
Allergic, neurasthenic, and dyspeptic symptoms, relations, 328
Alternation, individual differences, 334
Anthropology and the abormal, 59
Apparatus, flexible elevated finger maze, 482
  electric-current–stabilizing devices, 477
  electrical maze stylus, 485
  reflexohmeter, 238

Behavior and psychophysics, 299
Behaviorism and voluntary activity, 198

Caloric nystagmus in rabbit, 440
Color blends, imagined, 432
Conditioned reflex, explanation of habit formation, 110, 219
  temporal factors, 173
  and voluntary activity, 198

Dark-adaptation, duration and intensity of exposure light as factors, 4, 20
Difficulty as factor in effect of practice on individual differences, 249
Duration and intensity of exposure light and dark-adaptation, 4, 20
Dyspeptic, neurasthenic, and allergic symptoms, relations, 328

Emotion, specificity of facial expression, 42
Expression, facial, specificity, 42
Eyelid reaction, conditioning of, 173

Fields of instruction and research in psychology, 211
Franz, S. I., obituary, 3

Gestalt, structure of acts, 450

Habit formation and conditioned reflex, 110, 129

Illusions, effect of combining, 472
Imagined color blends, 432
Individual differences in ability to alternate, 334
  effect of practice with measurements weighted for difficulty, 249
Intelligence, influence of speed and age, 208
  and speed attitude, 465

Kinephantom and age, 204

Learning, practice and individual differences, 249
  retention and symbolic rehearsal, 376
    retroaction and recognition method, 234
    warming-up effect, 214
Luminal, effect on psychophysiological reactions, 415
Luminosity, algebra of, 286

Memory, effect of music on, 137
Music, effect on memory, 137

Neurasthenic, dyspeptic, and allergic symptoms, relations, 328
Nystagmus, caloric, in rabbit, 440

Perception of human figure, tachistoscopic, 152
Perseverative tendency in motor activity, 455
Phenobarbital, effect on psychophysiological reactions, 415
Practice and individual differences, 249
Psychology and techniques of discovery, 364
Psychophysics and behavior, 299
Psychophysiological measurements during phenobarbital somnolence, 415

Rabbit, caloric nystagmus in, 440
Recognition method and retroaction, 234
Reflexohmeter, 238
Retention and symbolic rehearsal, 376

Retroaction and method of recognition, 234

Size of pupil and form of color fields, 83
Space perception in translocated tissue, 227
Specificity of facial expression, 42
Speech, detection of tempo by visual, tactual, and visual-tactual cues, 100
  measurement of certain phases, 399
Speed attitude and test performance, 465
  and intelligence scores, 208
Stroboscopic effect, Metfessel's, explanation, 463
Structure of acts, 450
Summation in human retina, 311
Symbolic rehearsal and retention, 376

Tachistoscopic perception of human figure, 152
Tactile discrimination, improvement following localized training, 469
Tactual perception of space in translocated tissue, 227
  of speech, 100
Tempo, absolute, 230
Temporal factors in conditioning, 173
Test performance and speed attitude, 465

Vision, effect of combining illusions, 472
  of size of pupil on form and color fields, 83
  exposure light in dark-adaptation, 4, 20
  imagined color blends, 432
  spatial summation of subliminal stimuli in retina, 311
Voluntary activity and behaviorism, 198

Warming-up in a perceptual-motor task, 214